IN 1758

JPL PUBLICATION 81-95

Proceedings of the Seventeenth Annual Conference on Manual Control

University of California, Los Angeles, California
June 16-18, 1981

October 15, 1981

Prepared for
Office of Naval Research
Through an agreement with
National Aeronautics and Space Administration
by
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

JPL PUBLICATION 81-95

Proceedings of the Seventeenth Annual Conference on Manual Control

University of California, Los Angeles, California
June 16-18, 1981

Co-hosted by the

Jet Propulsion Laboratory and Industrial Associates of
California Institute of Technology
Pasadena, California

Co-sponsored by

NASA Headquarters
Air Force Flight Dynamics Laboratory
and
Office of Naval Research

Conference Co-chairmen

John Lyman
Biotechnology Laboratory
UCLA
Los Angeles, California 90024

Antal K. Bejczy
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109

October 15, 1981

Prepared for
Office of Naval Research
Through an agreement with
National Aeronautics and Space Administration
by
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Prepared by the Jet Propulsion Laboratory, California Institute of Technology, for the Office of Naval Research, through an agreement with the National Aeronautics and Space Administration.

FOREWORD

This volume, published with the support of the Engineering Psychology Programs of the Office of Naval Research MIPR No. N0001481MP0020, contains the proceedings of the Seventeenth Annual Conference on Manual Control held at the University of California, Los Angeles, CA from June 16 through 18, 1981. It contains complete manuscripts of most of the papers presented at the conference and abstracts of the others. Both formal and informal papers were presented, as indicated in the Table of Contents. The papers and abstracts are ordered as presented in the conference program, with session titles indicated in the Table of Contents. The papers and abstracts are reproduced in the proceedings from the original manuscript of the authors.

This was the seventeenth in a series of annual conferences on manual control dating back to December, 1964. These earlier meetings and their proceedings are listed below:

First Annual NASA-University Conference on Manual Control, The University of Michigan, December 1964. (Proceedings not printed.)

Second Annual NASA-University Conference on Manual Control, Massachusetts Institute of Technology, February 28 to March 2, 1966, NASA SP-128.

Third Annual NASA-University Conference on Manual Control, University of Southern California, March 1-3, 1967, NASA SP-144.

Fourth Annual NASA-University Conference on Manual Control, The University of Michigan, March 21-23, 1968, NASA SP-192.

Fifth Annual NASA-University Conference on Manual Control, Massachusetts Institute of Technology, March 27-29, 1969, NASA SP-215.

Sixth Annual Conference on Manual Control, Wright-Patterson AFB, April 7-9, 1970, proceedings published as AFIT/AFFDL Report, no number.

Seventh Annual Conference on Manual Control, University of Southern California, June 2-4, 1971, NASA SP-281.

Eighth Annual Conference on Manual Control, University of Michigan, Ann Arbor, Michigan, May 17-19, 1978, proceedings published by MIT, no number.

Ninth Annual Conference on Manual Control, Massachusetts Institute of Technology, May 23-24, 1973, proceedings published by MIT, no number.

Tenth Annual Conference on Manual Control, Wright-Patterson AFB, April 9-11, 1974, proceedings published as AFIT/AFFDL Report, no number.

Eleventh Annual Conference on Manual Control, NASA-Ames Research Center, May 21-23, 1975, NASA TM X-62, 464.

Twelfth Annual Conference on Manual Control, University of Illinois, May 25-27, 1976, NASA TM X-73, 170.

Thirteenth Annual Conference on Manual Control, Massachusetts Institute of Technology, June 15-17, 1977, proceedings published by MIT, no number.

Fourteenth Annual Conference on Manual Control, University of Southern California, Los Angeles, CA, April 25-27, 1978, NASA CP-2060.

Fifteenth Annual Conference on Manual Control, Wright State University, March 20-22, 1979, AFFDL-TR-79-3134.

Sixteenth Annual Conference on Manual Control, Massachusetts Institute of Technology, May 5-7, 1980; proceedings published by MIT, no number.

The main theme of the Seventeenth Conference was: "Perceptive/Cognitive Man-Machine Interaction and Interface."

John Lyman, UCLA

Antal K. Bejczy, JPL-CALTECH

TABLE OF CONTENTS

*F: Formal Paper
*I: Informal Paper

TABLE OF CONTENTS (Continued) Page

TABLE OF CONTENTS (Continued) Page

TABLE OF CONTENTS (Continued) Page

TABLE OF CONTENTS (Continued) Page

TABLE OF CONTENTS (Continued) Page

TABLE OF CONTENTS (Continued) Page

PUPILLOMETRIC MEASUREMENT OF OPERATOR WORKLOAD

Jackson Beatty
Department of Psychology
and
Brain Research Institute
University of California, Los Angeles

Operator workload and its assessment is a major issue in human performance theory. It is intuitively compelling to hold that some mental operations demand more of the operator than do others. We speak easily of attention-demanding tasks and tasks that can be performed "in one's sleep." Yet we know very little about what these demands might be, what the structure or structures are upon which such demands are placed, and, not surprisingly, we have reached little concensus as to how processing load might be measured. In this paper, I shall describe one approach to the workload measurement, pupillometry. Pupillometric measures provide an indication of momentary fluctuations in central nervous system excitability that occur as cognitive operations are performed; the magnitude of these changes may serve as a sensitive indicator of the workload imposed by cognitive tasks.

PUPILLOMETRY

Pupillometry utilizes optical measurement methods to determine the diameter of the pupil of the eye. Photographic measurements were employed in early pupillometric experiments; now electronic video pupillometry is commonly used. These instruments, such as the G and W Applied Science Laboratories television pupillometer, process an infrared video image of the eye to extract either vertical pupillary diameter or, in some cases, pupil area. This value is recomputed 30 times each second. In simpler systems a headrest is used to stablize the position of the pupil in space; more complex systems allow free head movement and track the position of the eye using a second, larger video image to control a servomechanism that aims the pupil video camera. Such instruments have been installed recently in aircraft cockpit simulators.

In our laboratory, the output of the pupillometer is sampled and stored by computer for later analysis. Three standard programs are utilized: the first performs an inspection of the data, correcting trials with minor artifacts and rejecting trials with more seriously contaminated data; the second program averages trials together, sorting by stimulus and response codes as appropriate; the third computes the changes in pupillary diameter over specified segments of the trial from

the averaged task-evoked pupillary responses. Similar procedures are employed in other laboratories.

PHYSIOLOGICAL CONTROL OF PUPILLARY MOVEMENTS

Pupillary diameter is determined by the relative strengths of contraction of the two opposing muscle groups of the iris, the dilator and sphincter pupillae. The dilator muscles are radially oriented bands of smooth muscle that are innervated by the sympathetic branch of the autonomic nervous system. Contraction of these muscles dilates the pupil. Conversely, the sphincter pupillae are innervated by the parasympathetic system and act to close the pupil when activated. Pupillary dilation, therefore, can result from either sympathetic excitation or parasympathetic inhibition. Both these pathways are affected by activation of nuclei comprising the reticular activating system of the brainstem; thus, pupillary movements are used in classical neurophysiology to measure the activation of these reticular nuclei (Moruzzi, 1972).

Although the details of the interaction of the reticular core with higher brain regions are only poorly understood at present (Hobson and Brazier, 1980), there is little doubt that cortico-reticular and reticulo-cortical interactions play a major role in determining the dynamics of higher brain functions. Most commonly the reticular system is viewed as serving an energizing function for the cortex (Luria, 1973), but it is perhaps wise to regard this viewpoint as a metaphor. Nonetheless, the ideas that pupillary movements reflect reticular activation and that reticular activation controls the dynamics of cognitive processing provide a theoretical basis for the use of pupillometry in the study of mental workload.

WITHIN-TASK VARIATIONS IN PROCESSING LOAD

The amplitude of the task-evoked pupillary response has been shown to reflect variations of processing load within a wide variety of cognitive tasks. Some examples are the following.

Memory. One of the first clear demonstrations that pupillary diameter varys with processing load was found in the study of short-term memory (Kahneman and Beatty, 1966; Kahneman, Beatty and Pollack, 1967). In simple short-term recall of digit strings, as when being told a telephone number and then dialing it, a characteristic pattern of pupillary movements are observed. Pupil diameter increases systematically as each digit is heard, reaching a maximum dilation in the interval before report (subjects in these experiments repeated the digit strings at the rate of 1/sec rather than dialing a telephone). Furthermore, the size of the peak dilation between listening and report is a monotonic function of the number of items heard (3 to 7 digits). During report, the pupil constricts with each digit spoken, reaching baseline diameter at the completion of the task. Such effects are highly reliable; they occur in every subject and are remarkably

consistent in magnitude. The dilation for 7 digits is approximately 0.5 mm.

The slope of these task-evoked pupillary responses is determined by the difficulty of the to-be-remembered information. For example, the memory span for unrelated nouns is shorter than that for digits; it is said, therefore, that the nouns place larger demands on the processing system than do digits. Thus, it is not surprising that the magnitude of pupillary dilation for individual items is greater for nouns than digits (Kahneman and Beatty, 1966).

A third important feature of the task-evoked pupillary response in the short-term memory task is that it increases with items presented only within the range of possible performance; when further demands are placed upon the processing system no additional dilation is observed. This important finding was first reported by Peavler (1974), who measured the task-evoked pupillary response for string lengths greater than the memory span. For most individuals, by the way, the span for digits is about 7 items. When longer strings are presented, errors begin to occur (Miller, 1956). Peavler found that when superspan strings were presented, the pupil dilated with each item until the seventh; subsequent items elicited no further dilation. This saturation of the pupillary response at the limits of capacity has been verified subsequently by other workers.

The short-term memory task provides a particularly clear demonstration of the properties of the pupillary response as a measure of mental workload. Manipulations that should increase load, here item type and number of items, increase the amplitude of the response in an orderly manner. Increases in processing demands beyond the capacity of the system to respond are not reflected in pupillary movements. The task-evoked pupillary response indexes the processing demands being met by the system, not the demands placed upon the system by the task. Thus, it may serve as a measure of mental work executed, at least within the confines of a particular task.

Language. A number of levels of language processing have been studied pupillometrically. Beatty and Wagoner (1978) showed that the degree of processing required in a simple letter matching task was faithfully reflected in the magnitude of the task-evoked pupillary response. The smallest dilations occured when only matching of physical details was require to reach a judgement of "same", as for the letter pair "AA". When name code extraction was required (as for "Aa") the pupillary response was significantly enhanced. But in both cases the response was relatively small, on the order of 0.1 mm.

Larger responses were observed by Ahern and Beatty (1981) in a word matching task, in which individuals heard two words and were required to determine if they shared the same meaning. The word pairs, which were drawn from psychometric tests of intelligence, varied in difficulty. Difficulty level was reflected in the amplitude of the task-evoked pupillary responses.

The largest responses for language processing were obtained with a grammatical reasoning task (Baddeley, 1968). Subjects listened to sentences of the form "A follows (precedes) B" and then heard an exemplar letter pair ("BA"); the task was to determine if the sentence correctly describes the pair. The sentences varied in grammatical form, being active or passive and positive or negated. A significant effect of grammatical complexity was observed. Response were approximately 0.5 mm. in amplitude (Ahern and Beatty, 1981).

Other tasks. Pupillometric investigations have also been reported for a variety of other types of tasks, including simple sensory and motor tasks. Although space does not permit a summary of these results, they support the conclusion that the task-evoked pupillary response is a sensitive and accurate physiological measure of within task variations in processing load.

BETWEEN-TASK VARIATIONS IN PROCESSING LOAD

But any useful measure of operator workload must do more than reflect within-task variations in processing load; it must, in addition, provide a measure of workload imposed by qualitatively different mental operations. Only in this way can operator workload be assessed in complex man/machine systems.

Evidence that the task-evoked pupillary response is sensitive to variations in processing load imposed by qualitatively different mental operations has been provided in a detailed review by Beatty (in press). From the published literature, all data were employed that met two criteria; there were no motor responses occuring during measurement and the published figures permitted estimation of the maximum value of the pupillary response. The resulting data were remarkably consistent, leading to the conclusion that the task-evoked pupillary response provides a reliable and reasonable indication of the processing load imposed by cognitive tasks that differ markedly in their internal structures.

BETWEEN-INDIVIDUAL VARIATIONS IN PROCESSING LOAD

Finally, there is evidence that the amplitude of the task-evoked pupillary response provides an indication of differences in processing load imposed by the same task on individuals who differ in cognitive abilities. Ahern and Beatty (1979, 1981), for example, measured pupillary responses while solving multiplication problems in university student with high and low Scholastic Aptitude Test scores. The high-scoring students showed significantly smaller pupillary responses at all difficulty levels. Further, the amplitude of these responses correlated only with ability measures and not with personality variables or psychometric measures of state or trait anxiety. Finally, there was no difference between groups in the amplitude of the light and dark pupillary reflexes. These data support the view that solving the same objective multiplication problem (e.g., 6 time 8) is a more demanding

task for the less able students.

SUMMARY

All available evidence supports the view that the amplitude of the task-evoked pupillary response provides a sensitive indicator of the workload imposed by mental operations. However, much remains to be learned. We do not understand with any certainty the physiological mechanisms linking the autonomic periphery with the highest levels of central nervous system function. We know little about the ways in which these responses, or other similar physiological measures, may be applied to the solution of practical design problems facing human factors engineering. But the future appears promising; pupillary movements may provide one key for providing a solid empirical basis for the problem of operator workload assessment.

REFERENCES

Ahern, S.K., and Beatty, J. Physiological signs of information processing vary with intelligence. Science, 1979, 205, 1289-1292.

Ahern, S.K., and Beatty, J. Physiological evidence that demand for processing capacity varies with intelligence. In M. Friedman, J.P. Dos and N. O'Connor (Eds.), Intelligence and learning. New York: Plenum, 1981.

Baddeley, A.D. A three-minute reasoning test based on grammatical transformation. Psychonomic Science, 1968, 10, 341-342.

Beatty, J. Task-evoked pupillary responses, processing load, and the structure of processing resources. Psychological Bulletin, in press.

Beatty, J., and Wagoner, B.L. Pupillometric signs of brain activation vary with level of cognitive processing. Science, 1978, 199, 1216-1218.

Hobson, J.A., and Brazier, M.A.B. (Eds.), The reticular formation revisited: Specifying function for a nonspecific system. New York: Raven, 1980.

Kahneman, D., and Beatty, J. Pupil diameter and load on memory. Science, 1966, 154, 1583-1585.

Kahneman, D., Beatty, J., and Pollack, I. Perceptual deficit during a mental task. Science, 1967, 157, 218-219.

Luria, A.R. The working brain. New York: Basic Books, 1973.

Miller, G.A. The magical number seven, plus or minus two: Some

limits of our capacity for processing information. Psychological Review, 1956, 63, 81-97.

Moruzzi, G. The sleep-waking cycle. Reviews of Physiology: Biochemistry and Experimental Pharmacology, New York: Springer-Verlag, 1972.

Peavler, W.S. Individual differences in pupil size and performance. In M. Janisse (Ed.), Pupillary dynamics and behavior. New York: Plenum, 1974.

NOTE

This research was sponsored by the Environmental Physiology Program of the Office of Naval Research under Contract N00014-76-C-0616.

D2

THE COGNITIVE DEMANDS OF SECOND ORDER MANUAL CONTROL: APPLICATIONS OF THE EVENT RELATED BRAIN POTENTIAL[1]

Christopher Wickens, Richard Gill, Arthur Kramer, William Ross & Emanuel Donchin
University of Illinois Department of Psychology
Champaign, Illinois 61820

ABSTRACT

Three experiments are described in which tracking difficulty is varied in the presence of a covert tone discrimination task. Event related brain potentials (ERPs) elicited by the tones are employed as an index of the resource demands of tracking. Experiments 1 & 2 varied tracking difficulty by the order of the control system dynamics in a teleoperator target acquisition task (experiment 1), and in random input compensatory tracking (experiment 2). The ERP measure reflected the control order variation, and this variable was thereby assumed to compete for perceptual/central processing resources. A fine-grained analysis of the results of Experiment 2 suggested that the primary demands of second order tracking involve the central processing operations of maintaining a more complex internal model of the dynamic system, rather than the perceptual demands of higher derivative perception. Experiment 3 varied tracking bandwidth in random input tracking, and the ERP was unaffected. Bandwidth was then inferred to compete for response-related processing resources that are independent of the ERP.

[1]Experiment 1 was supported by a subcontract from NASA-Jet Propulsion Laboratory (US NASA SBC CAL TECH 955610). Dr. John Hestenes was technical monitor. Experiments 2 and 3 were supported by a contract from the Air Force Office of Scientific Research Life Sciences Directorate (F49620-79-C-0233). Dr. Alfred Fregly was the technical monitor.

Two variables that are well established to influence the difficulty of manual control are the order of the controlled system, and the bandwidth of a disturbing noise function (McRuer & Jex, 1968). As shown in the data of figure 1, both variables influence tracking error, as well as the perceived difficulty of the task.[1] Furthermore, figure 1 suggests that for each variable, both subjective and performance measures increase linearly. An approximate level of bandwidth can be established that produces roughly equivalent difficulty ratings and performance to second order tracking (.55-.60 Hz).

A change in task parameters, such as bandwidth or system order, that attenuates performance and increases subjective effort must be assumed to consume a greater quantity of the operator's limited capacity processing resources. However, given recent experimental evidence that these resources do not reside within a single undifferentiated "reservoir" (Navon and Gopher, 1979, Wickens, 1980), but are instead multidimensional, defined by two processing stages (perceptual-central vs response) and two processing codes (spatial vs verbal), then the resource demands of a task must be a vector quantity. It is not, however, immediately clear how the vector of demand increase imposed by order and bandwidth is defined.

In the case of control order, the increased demand imposed by second order control could be perceptual, given the requirement to generate lead and respond to higher derivatives of the error signal. It could be central, given the more complex internal model defined by two state variables rather than one, or it could be response, given the "bang-bang" or double impulse control sometimes employed in the regulation of higher order systems. In the case of bandwidth, an increase in this parameter will influence the frequency with which information must be sampled (perception), and corrective responses must be selected (response). It is reasonable to assume that in both manipulations the increased load is imposed upon spatial processing more than upon verbal, given the fundimental spatial nature of the manual control task.

In order to assess the stage-defined locus of these variables, it is necessary to employ a workload measure that is selectively sensitive to early or late processing stages. Such a measure is provided by the late positive component of the event-related brain potential, or ERP. The ERP is a transient series of voltage oscillations recorded by electrodes from the surface of the scalp and elicited in response to discrete environmental or cognitive events. The late positive, or P300, component is a deflection in the ERP that appears to be particularly sensitive to the attention or processing resources allocated to the eliciting stimulus (Donchin, 1980). In previous investigations, we have shown that the P300, elicited by counted stimuli in a Bernoulli series, is particularly sensitive to the perceptual load imposed by a concurrent display monitoring task (Isreal, Wickens, Chesney, & Donchin, 1980), but is unaffected by the response load of a manual control task (Isreal, Chesney, Wickens, &

[1]The 1.5 order point is a condition in which the output of a first and second order system are combined with equal weighting.

Donchin, 1980). Assuming that P300 acts as a selective measure of perceptual/central processing load, the present series of experiments examines the sensitivity of this measure to increases in control order and in bandwidth, in order to infer the stage-defined locus of these manipulations.

Experiment 1: Teleoperator Simulation

Subjects in this experiment engaged in a target acquisition and capture task, depicted in figure 2. A target appeared at a random location and traversed the display in a linear path. The subjects' task was to manipulate the cursor, via the two axis control in such a way as to match the spatial velocity of the target (acquisition phase). Control dynamics could be set at either first or second order. Once acquisition was completed, the target began to tumble, and the subject then had to match the angular velocity as well by a left hand control (orientation phase). When error in all three dimensions was held below a certain criterion, a successful capture could be affected by depressing a "capture" button with the thumb of the left hand. As the task was performed, either the target or the cursor stimulus randomly intensified or "flashed." In the data to be reported, the subjects' task was to maintain a covert count of the number of times the target intensified during a trial. The P300 component of the ERPs generated by these intensifications served as the workload index.

Figure 3aindicates the P300 (the large downward deflection) elicited during the acquisition (top) and orientation (bottom) phase. Clearly the latter phase imposes greater workload (3 axes of control as opposed to two) and this is indicated by the diminished P300s elicited by the relevant counted targets. (The non-counted cursor flashes failed to differentiate workload in any condition.) Of particular interest to the current hypothesis is the diminished P300 during second, as opposed to first, order control in both phases. These results support the hypothesis that a major source of workload in the higher order control condition is the increased perceptual load imposed by the perceptual anticipation, or central processing load, imposed by internal model updating.

In a second experiment, when subjects were then provided extensive practice with both control dynamics of the target acquisition task, the difference in P300 between first and second order tracking was eliminated (although acquisition performance in the former continued to be superior). Our interpretation of these results (figure 3b) in light of the selective nature of P300 and the totally predictable linear trajectory of the target path, is that with extensive practice, subjects in second order tracking employ double-impulse control strategy to place the target on the right trajectory, and need not engage in further load inducing perceptual/anticipatory processing.

Experiment 2: Steady State Compensatory Tracking

Unlike experiment 1, subjects in experiment 2 engaged in three-minute trials of compensatory tracking of a random disturbance input with an upper

cutoff frequency of 0.30 Hz. Again tracking order (first vs second) was the manipulated variable. A Bernoulli series of auditory tones, of high and low pitch occurring at an interval of 1.5 seconds were the stimuli enployed to generate the ERPs. Figure 4 shows the averaged ERPs elicited by the counted auditory tones when the subject was not tracking at all (solid line), was tracking a first order system (dashed line) and was tracking a second order system (dotted line). In agreement with experiment 1, these results manifest the same attenuation of P300 with order ($p < .05$). In addition, they suggest the large differences in P300 between the condition of focussed attention and that when resources are diverted to tracking ($p < .01$).

Experiment 3: The Effect of Bandwidth

The same subjects that participated in experiment 2 were employed in an investigation of the bandwidth variable. In this study, employing only first order dynamics, a careful calibration was first performed to determine for each subject the level of bandwidth that would generate equivalent error to that observed in second order tracking. We assume, on the basis of the data from figure 1, that manipulations of order and bandwidth that generate equivalent changes in error, will also generate equivalent changes in subjective difficulty. Thus by the calibration process employed in experiment 3, we have made the bandwidth manipulation as equivalent as possible to the order manipulation of experiment 2.

The ERPs,generated and recorded exactly as in the previous experiments, are shown in figure 5. Again a marked decrease in P300 amplitude is observed between the control condition and "easy" tracking. However, no further attenuation is observed between low and high bandwidth tracking. In fact, there appears to be a slight increase in amplitude. These results represent a replication of the finding of Isreal, Chesney, Wickens, and Donchin (1980) in which they observed that P300 did not diminish as bandwidth was gradually increased from a low to high level. The results in the present case are convincing because of the comparability of the bandwidth with the order manipulation in terms of tracking error and subjective difficulty. Order attenuated P300, while bandwidth did not.

Discussion

The localization of demands of the tracking order manipulations at earlier processing stages is in agreement with another study of Wickens and Derrick (1981), in which the Sternberg Memory Search Task was employed as a selective workload index. Furthermore, this investigation failed to find any evidence for an interaction of control order with response-load variables of the Sternberg Task. Thus control order was not assumed to be response loading in their task. Discriminating between perceptual and central processing loci of second order control is somewhat difficult on the basis of the task interference data discussed above, since the multiple resources model postulated by Wickens (1980) assumes that both perceptual and central processing operations compete for common resources, and therefore will not differentially interfere with a secondary task.

To discriminate the perceptual from the central locus, the following more detailed analysis was thus undertaken. It was assumed that if the resource demands of second order tracking were directly attributable to the perception of higher error derivatives, then these demands should covary with the momentary state of the system. When error acceleration is high, demands should increase and P300s would be attenuated;when low,demands should decrease, and P300 enhanced. In contrast, if the demands were related to central processing, we assume that the more complex internal model of a second order system is activated at the beginning of the trial, and the resources required to monitor this activation are relatively continuous, and independent of momentary fluctuations in system state.

To contrast these hypotheses, P300s were selectively averaged according to the momentary state of the system at the time of stimulus presentation. "State" was defined along six dimensions (i.e., whether error position, velocity, or acceleration was above or below its median value, and whether stick position, velocity, or acceleration was above or below the median). In general, the results supported the central-processing hypothesis. With only minimal exceptions, all subjects seemed to show equivalent P300s at high and low ends of the averaging categories, no matter how these categories were defined (i.e., by e, ė, ë; s, ṡ, or s̈). Interestingly, this equivalence was observed whether subjects were good or poor trackers. We are not surprised to observe the independence from the control stick variable, given the assumed independence of P300 from response variables. The independence from variation in perceptual (error) state suggests that the increased resource demands of second order control are relatively continuous, and therefore likely to be of central origin.

A second aspect of the results concerns the locus of the bandwidth manipulation. It is obvious that increasing the bandwidth does indeed increase the required frequency of perceptual sampling. However, as suggested by the data of Wickens and Derrick (1981), or Isreal, Chesney, Wickens & Donchin (1980), and experiment 3, this increase is not of sufficient magnitude to compete with a second task for resources. By process of elimination, the locus of bandwidth demands would seemingly be response related, although this hypothesis has yet to be explicitly tested. The failure of P300 to reflect bandwidth variation is important because it emphasizes the selective aspect of the ERP measurement and the multidimensionality of processing resources. Were P300 affect by any manipulation of task difficulty, then the present results would contribute little in the way of added knowledge. However, the selectivity observed suggests that different variations of manual control difficulty exert a qualitative as well a quantitative influence on the resources of the human operator.

References

Donchin, E. Event-related brain potentials: A tool in the study of human information processing. In H. Begleiter (Ed.), Evoked potentials and behavior. New York: Plenum Press, 1979, pp. 13-75.

Isreal, J.B., Chesney, G.L., Wickens, C.D., & Donchin, E. P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. Psychophysiology, 1980, 17, 259-273.

Isreal, J.B., Wickens, C.D., Chesney, G.L., & Donchin, E. The event-related brain potential as an index of display-monitoring workload. Human Factors, 1980, 22(2), 212-224.

McRuer, D., & Jex, H. A review of quasi-linear pilot models. IEEE Translations on Human Factors in Electronics, 1967 HFE-8, 231-249.

Navon, D. & Gopher, D. On the economy of the human processing system. Psychological Review, 1979,

Wickens, C.D. The structure of attentional resources. In R. Nickerson (Ed.), Attention and Performance VIII. Erlbaum, 1980.

Wickens, C.D. & Derrick, W. The processing demands of higher order manual control: Application of additive factors methodology. University of Illinois Engineering Psychology Laboratory Technical Report EPL-81-1/ ONR-81-1, March, 1981.

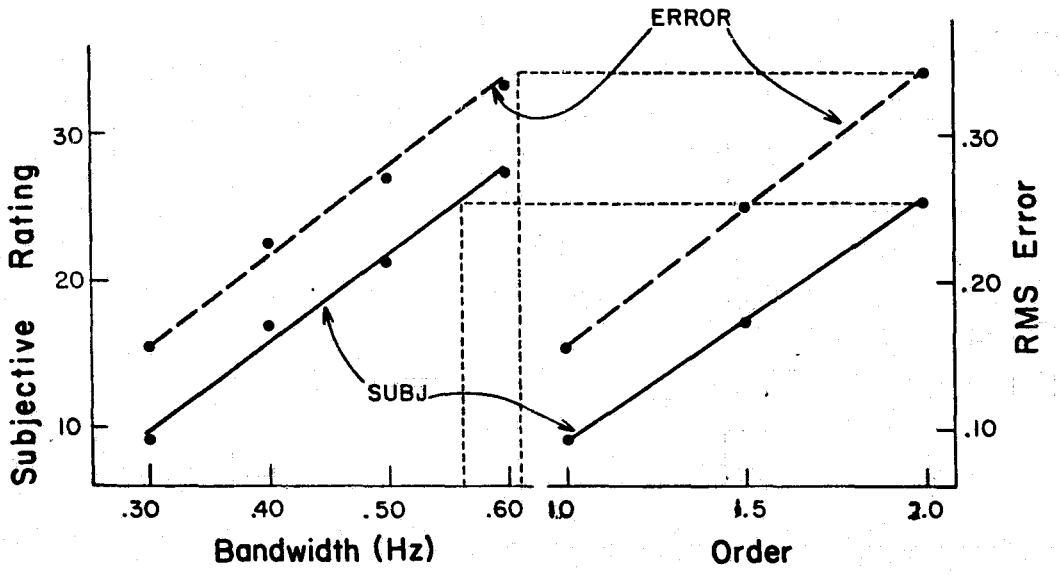


Figure 1.

Influence of bandwidth and control order on tracking error and subjective ratings of difficulty.

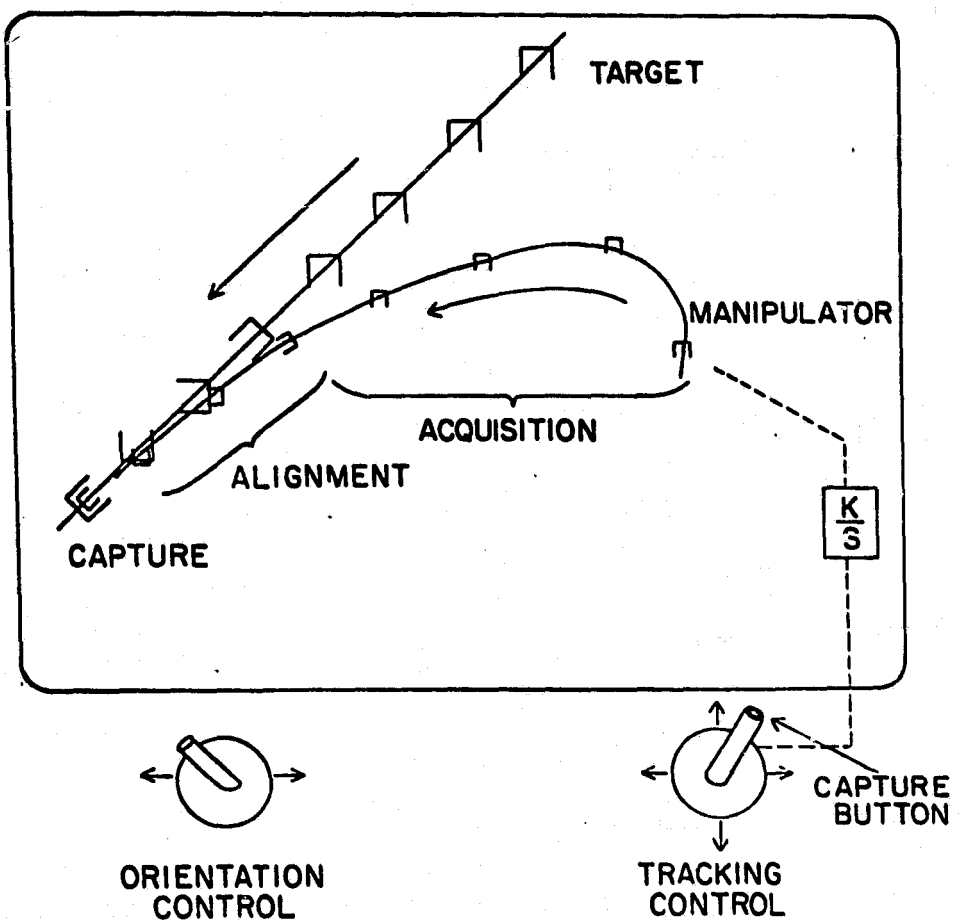


Figure 2.

Experimental paradigm for teleoperator target acquisition task.

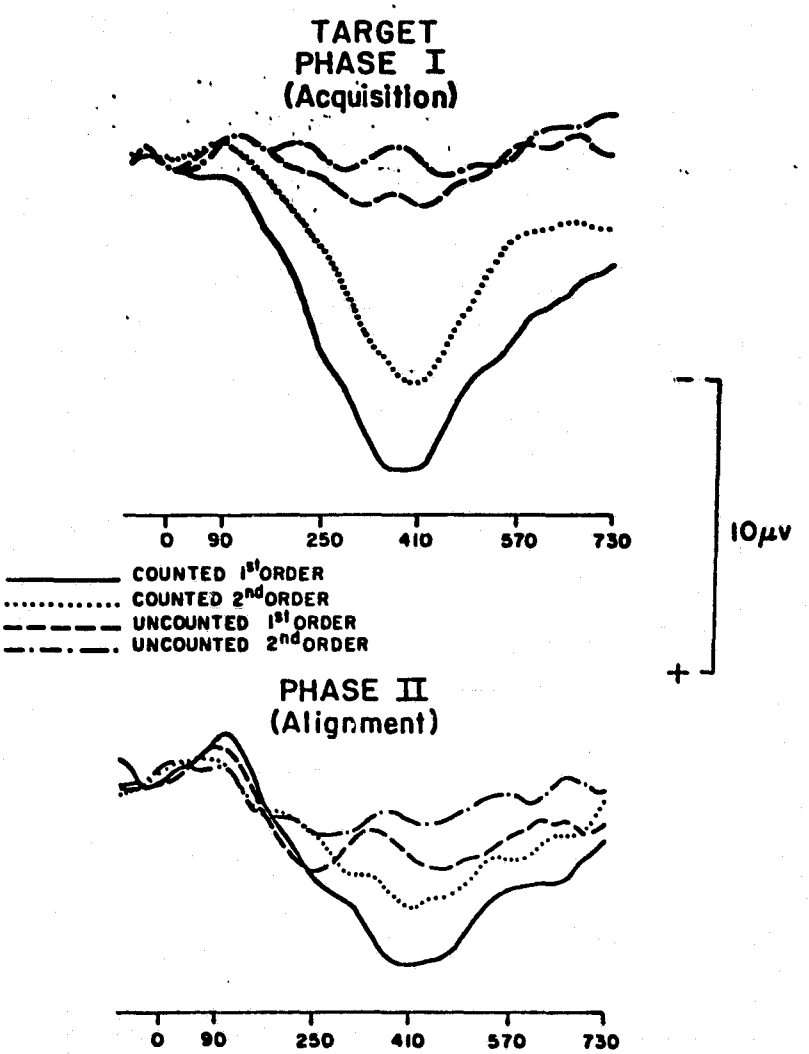


Figure 3(a)

(Experiment 1: Target Acquisition)
Effect of control order on the ERP.

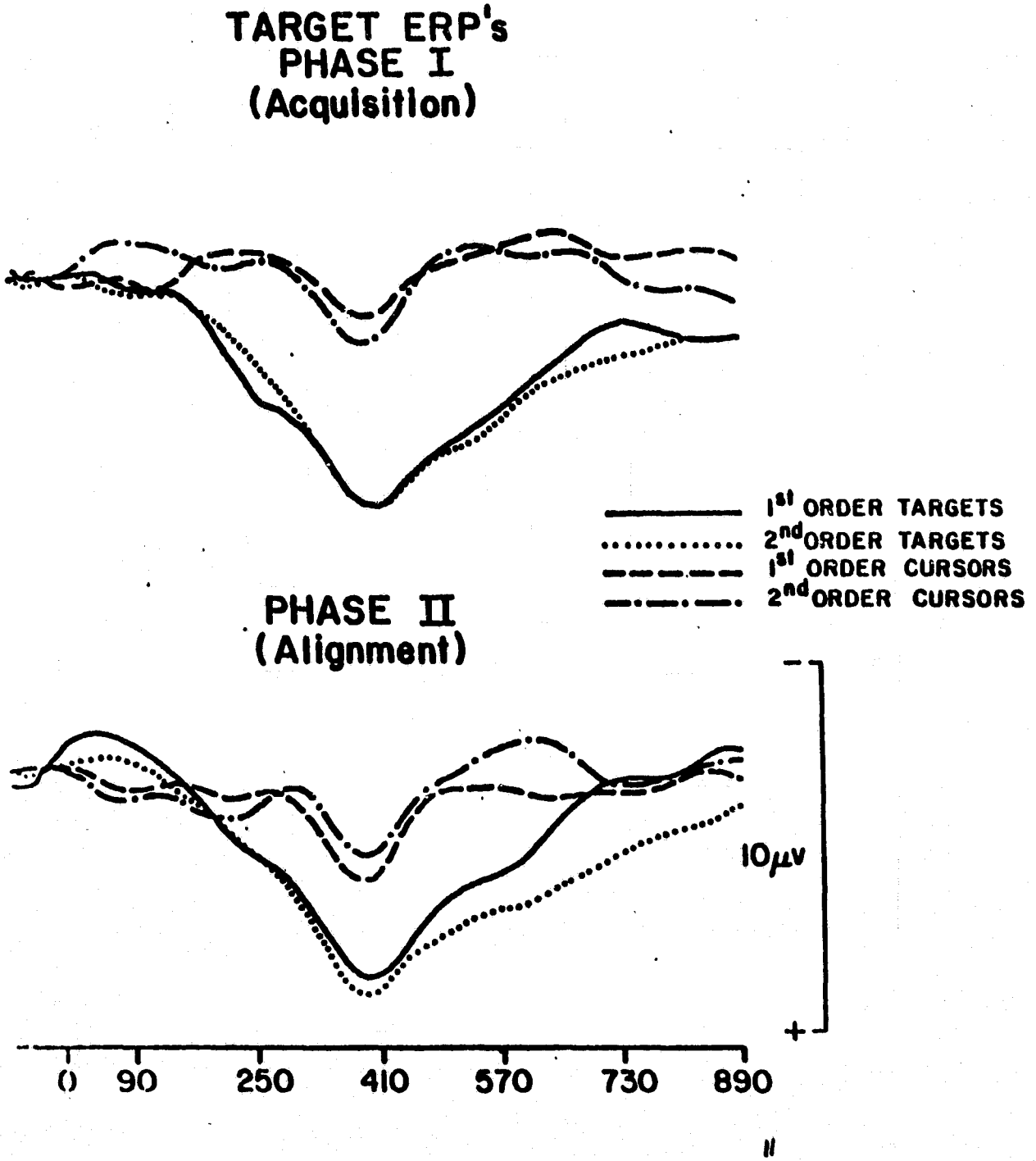


Figure 3(b)

(Experiment 1: Target Acquisition)
Highly practiced subjects (570 acquisition trials).

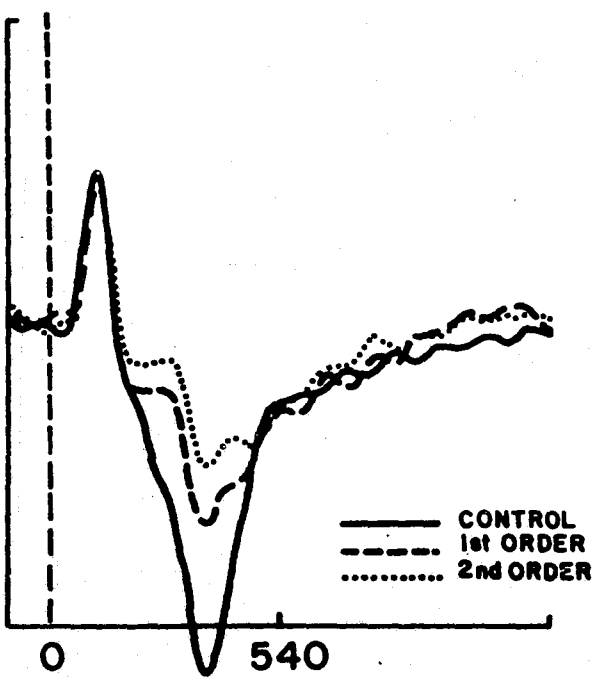


Figure 4.

Experiment 2: Random input tracking: Effect of control order on the ERP.

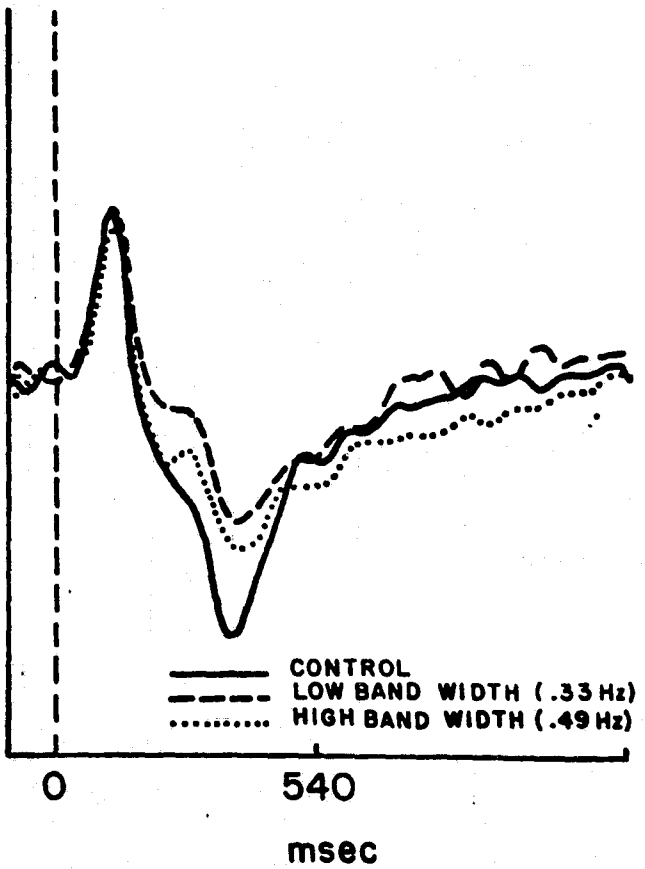


Figure 5.

Experiment 3: Random input tracking: Effect of bandwidth on the ERP.

D3

A Study on Task Difficulty and Acceleration Stress

D. W. Repperger D. B. Rogers

Air Force Aerospace Medical Research Laboratory
Wright Patterson Air Force Base, Ohio 45433

SUMMARY

The results of two experiments are discussed which relate to task difficulty and the effects of environmental stress on tracking performance. The first experiment involved 5 different sum of sine tracking tasks which humans tracked both in a static condition and under a 5 Gz acceleration stress condition. The tasks were designed in such a manner as to investigate workload measures and to compare our hypothetical design to subjective evaluations. The tasks were required to satisfy 5 criteria specified in mathematical terms.

The second experiment involved similar environmental stress conditions but in this case the tasks were constructed from deterministic functions with specially designed velocity and acceleration profiles. In both parts of this experiment, subjective evaluations were obtained and compared to the assumption that difficulty is related to magnitudes of velocity and acceleration profiles of the target tracking task. Phase Plane performance analysis was conducted across 7 subjects to study potential measures of workload or tracking difficulty.

INTRODUCTION

In the study of manual control theory, the systematic characterization of task difficulty has been a problem of considerable interest for many years. An extensive amount of work has been done in this area and a variety of studies indicating different measures related to workload are available in the Human Factors and Psychological literature. In the engineering literature, the classical paper by Cooper (reference 1) illustrates the motivation for such a characterization of task difficulty - a subjective rating scale for human tracking. The extent at which this subjective rating scale can be used to elict pilot response is best illustrated in reference 2 where a thorough study has been done to investigate and pinpoint the exact cause-effect relationships between pilot subjective ratings and handling qualities of aircraft. This study uses a decision tree type of analysis procedure to investigate the responses.

At the Air Force Aerospace Medical Research Laboratory, it is of interest in our research program to develop standard tasks or levels of tracking difficulty and to be able to estimate levels of difficulty associated with human tracking. Once a consistent set of tasks are developed which provide a basis or standard for tracking behavior, it is then possible to more closely evaluate the effects of stress on human tracking performance. The criteria for the design of the tracking tasks must be such that each task is required to be a sensitive indicator of performance change (between each task number) and, in addition, the requirement is made that the task is to be sufficiently sensitive as to

show a performance decrement between the stress-non stress condition.

This study consisted of two separate experiments. Both parts of this investigation involved human tracking for target forcing functions with different acceleration and velocity profiles. It was desired to study a critical task concept (reference 3) based on a hierarchy of difficulty associated with the different target forcing functions. This approach differs from the classical critical task concept considered by Jex, et al. (reference 4) in which the controlled element would have dynamics that change. In our studies, the controlled element (figure (1)) remained the same; the tracking tasks varied in levels of difficulty based on our hypothesis of different velocity and acceleration profiles associated with each target forcing function. The motivation for this work was due to an interesting paper by Verplank (reference 5) in which he equated difficulty and stress in studying human response behavior within a vigilance paradigm.

SYMBOLS

$f(t)$ = The Target Forcing Function Signal
$e(t)$ = The Closed Loop Error Signal
$x(t)$ = The Output of The Plant (Controlled Element)
R = Radius in the $\dot{f}$ versus $\ddot{f}$ plane = $(\dot{f})^2 + (\ddot{f})^2$
μ = Median of the distribution of the error window histogram
e_{RMS} = Root Mean Square error score
$\bar{x}$ = mean e_{RMS} value
σ = standard deviation of e_{RMS} value
x_d = The deviation of a difference from the mean of the differences.
M_d = Mean of the n differences of paired observations.
p = Probability
t = t test statistic
t = time

METHOD

Subjects - Seven male United States Air Force volunteers participated in this experiment. They had prior training in both the G type of stress exposures and manual tracking tasks.

Design of The Target Tracking Task - Part I

The objective of this study was to develop the tracking tasks of different levels of difficulty and to study their ability to produce performance decrements between tasks (for a given experimental condition) and between experimental conditions (for the same task). For the first part of this study, it was decided to design five different tasks with the following constraints:
(1) Each tracking task will be zero mean, constant variance, sum of sines.
(2) Each forcing function when presented as a replication will have a random

initial phase angle for each frequency component. The phase angle must be a prime multiple of the fundamental frequency and not a linear constant multiple of any other frequency component.
(3) Each forcing function will have a random phase angle between each frequency.
(4) Due to human physiological exposure limits in the design of the acceleration experiment, the length of each task was set at 15 seconds.
(5) The amplitudes of all the sinusoids are scaled so such that they all have equal power and produce the same displacement on the CRT (display). The open loop and autopilot runs of this study which verify this fact are presented in the sequel.
(6) The component frequencies of the sinuosids are "relatively prime" multiples of a fundamental frequency.
(7) A shift in frequency content is required so that $ff_i > ff_j$ is true if $i>j$, $i,j=1,2,..5$ where the frequency content of ff_i is higher than that of ff_j. The procedure for obtaining this desired result is discussed in reference 6.

Using a measure of difficulty denoted as R (the distance in the target phase plane (figure (2))) where R satisfies:

$$R^2 = (\dot{f})^2 + (\ddot{f})^2 \qquad (1)$$

Then table I illustrates the values of R obtained for the 5 different tasks chosen in Part I of this study.

Table I Results of The Open Loop and Autopilot Simulations

Forcing Function (or Task) Number	Open Loop Error RMS * 2351.	Autopilot Error RMS * 2351	R^2(mean) For The Forcing Function	R^2(s.d.) For The Forcing Function
#1	718.6	477.8	0.352	0.290
#2	718.6	477.8	0.662	0.611
#3	718.6	477.8	1.212	1.140
#4	718.6	477.8	2.151	1.897
#5	718.6	477.8	3.509	3.2309

In this table the results of the open loop and autopilot runs are also displayed. These results (columns 2 and 3) illustrate the consistency of the normality conditions imposed in this study on the task numbers.

Design of The Tracking Task - Part II

In this design, the object was to design a different type of target forcing funcion. Figure (3) illustrates the shape of the functions used in this part of the study. In this case the objective was to have forcing functions of varying difficulty. The assumption is that the radius R is a metric of dispersion about the origin defined by equation (1) and tasks with larger R values are more difficult to track. The design of the function in figure (3) is a result of the need to have target tracking tasks that varied the value of R as a function of time. To create the shape of the diagram in figure (3), the

following exponential functions were chosen based on a set of time intervals:

Time Interval(seconds)	Function Chosen	
$(t_1, t_2) = (0,5)$	$f_1(t) = \int_0^t a\, e^{-(s-2.5)^2/2\sigma^2} ds$	(2a)
$(t_2, t_3) = (5,10)$	$\dot{f}_2(x) = \dot{f}_1(5) + \int_0^x (-ae^{-(s-7.5)^2/2\sigma^2}) ds$	(2b)
$(t_3, t_4) = (10,15)$	$f_3(y) = f_2(10) + \int_0^y (-ae^{-(z-12.5)^2/2\sigma^2}) dz$	(2c)
$(t_4, t_5) = (15,20)$	$\dot{f}_4(z) = \dot{f}_3(15) + \int_0^z (ae^{-(t-17.5)^2/2\sigma^2} dt)$	(2d)

Where:

$x = t-5$ (3a)

$y = t-10$ (3b)

$z = t-15$ (3c)

With some manipulation, the relationships (2a-d) and (3a-c) can be shown to produce the trajectories displayed in figure (3). The value a can be adjusted to sweep out a range of values. Table II illustrates the values chosen for part II of this study:

Table II - Forcing Function Design For Part II

FF #	a	b
1	0.1	.04
2	0.2	.08
3	0.3	.12
4	0.4	.16
5	0.5	.20

Randomization of The Presentation of The Tasks

Reference 6 describes the procedure chosen to ensure that the subjects would not know the order of presentation of the five different tasks at any time during the experiment. This was true for Parts I and II for both the static and stress portions of the experiment.

Apparatus

A 19-foot arm centrifuge (figure 4) was used to establish a 5 Gz stress condition for the subjects. In Air Force applications this acceleration force is in the z direction (down the spine of the subject) and is termed Gz. The centrifuge rotated at an angular speed of 27.5 RPM with the cab vectored at 78 degrees about a line in the z axis of the subject. The subjects wore standard Air Force helmets, gloves, and an Anti-G Suit with a G-valve. The Anti-G Suit-G-valve delivers a specific air pressure to the bladders of the Anti-G Suit.

Training Orientation and Data Exposures

During this training orientation, the subjects were required to asymptote to the five tracking tasks (performance training) and also to acclimate to the G stress (physiological adaptation). In the final design of this experiment, each day's run consisted of five component parts or phases (figure 5 illustrates one day's run for data collection). Phase I comprised of the presentation of the five tracking tasks in the static condition (no stress). Phase II consisted of the presentation of the five tracking tasks at an acceleration stress level of 5 Gz with a 20-second preliminary warm-up run at 4 Gz. After the five exposures at 5 Gz, the centrifuge was brought to a stationary position and the subject again performed five tracking tasks presented in random order in the static condition (Phase III). Phase IV of the daily run consisted of five tracking tasks presented in random order again under the five Gz stress as in Phase II. In Phase V of this experiment, the five tasks were presented in the static mode. Again, as with all the previous tasks, all forcing functions were presented in a random sequence. Four data days were collected after the subject progressed satisfactorily in the indoctrination period. During the data collection phase of the experiment, the subject never experienced more than 300 seconds per day of 5 Gz exposure nor more than two daily exposures per week. After the 4 data days were collected, a questionnaire was administered on the fifth day with the subject sitting in the centrifuge but with no machine motion. The questionnaires recorded subjective ratings of the task difficulty hierarchy.

Questionnaire

One definition of workload (reference 7), indicates that it is a function of increased performance requirements plus additional attention requirements. To get a true subjective evaluation, it was necessary to ask the subjects how they rated the tracking tasks. On the last day of the experiment the subjects were presented 25 tasks in random order. After the first task, each subject was asked to compare the task he was presently tracking with the previous one. The subject was asked whether the present task was more difficult, less difficult, the same, or not possible to rate. Thus the subject, was not knowledgeable as to which forcing function number was presented and would only give relative ratings between tasks.

RESULTS

CDF Performance Results From The Data

As discussed previously, after the 5-day indoctrination period the seven subjects tracked until they trained to an asymptotic level of performance for the different tracking tasks. One criterion used to define asymptotic performance is that on three consecutive days, the RMS performance scores do not decrease more than 5% on daily exposures of 25-50 static presentations of

the five random targets per day. After this level was reached, the subjects were assumed to be trained. In part I of this study, the first question to be asked concerns the adaptation of the subjects to learning the tracking task and acclimation to G levels.

To address the question of learning and adaptation to stress, a table based on error scores was constructed across all seven subjects and four replications of each stress condition. Table III illustrates these results:

Table III - Stress Data $\bar{x}/\sigma$ Ratios For 7 Subjects, 2 Replications/Day

	Day 1	Day 2	Day 3	Day 4
ff#1	1.9	5.2	2.9	4.6
ff#2	2.6	4.9	3.6	4.0
ff#3	5.3	4.5	13.1	10.6
ff#4	8.2	10.0	17.1	13.5
ff#5	10.1	5.9	24.3	20.4

If any trends did exist in the data runs, due either to further performance training (reduction of tracking error) or possibly to further acclimation to G stress (physiological adaptation), they would be shown by a gradual increase in the ratio $\bar{x}/\sigma$ across a row for a given forcing function number. Since there appear to be no apparent trends for this stress acclimation, it is assumed that the subjects had adjusted to a steady state physiological conditioning and tracking performance level.

The next question to be addressed here is whether the forcing function number was correlated with measures of performance degradation. From the CDF figures (similar to figure 6), using data from all seven subjects (five replications), it was desired to conduct tests to investigate if a significant performance decrement exists dependent on forcing function number for both the static or the stress conditions. The following statistical test would determine this effect:

$$H_o: \quad \mu_{i+1} > \mu_i \qquad i=1,\ldots 4 \qquad (4a)$$

versus

$$H_1: \quad \mu_{i+1} \leq \mu_i \qquad i=1,\ldots 4 \qquad (4b)$$

where μ corresponds to the 0.5 line on the CDF in figure 6. This figure is illustrated here to show how the median point μ is obtained. This corresponds to a "median window" size for the tracking error signal. The test was conducted for both the static data and the stress data. The results using a t statistic are displayed in Table IV:

Table IV

Hypothesis test on values	$\bar{t}$ for static data	p	$\bar{t}$ for stress data	p
$ff_2 > ff_1$	5.174	<.01	11.48	<.01
$ff_3 > ff_2$	18.39	<.01	19.91	<.01
$ff_4 > ff_3$	7.15	<.01	9.64	<.01
$ff_5 > ff_4$	4.55	<.01	5.04	<.01

The $\bar{t}$ values used in Table IV were the t statistic (2-tailed test) for correlated data (references 8 and 9) which satisfies:

$$\bar{t} = Md \Big/ \sqrt{\frac{\sum_{i=1}^{n} x^2_d}{n(n-1)}} \qquad (5)$$

where Md = the mean of the n differences of paired observations and x_d = the deviation of a difference from the mean of the differences. This test is for paired samples; they are not independent but are correlated due to the five replications involved with all seven subjects in the static case, and four replications with the stress data. The results from Table IV indicate performance decrements correlated with the forcing function number. The performance decrement is significant at an .01 level for an increase of forcing function number in both the static and stressed condition. The test given here corresponds to changes in medians (i.e., the 0.5 point on the CDF curve). Using the CDF method, this analysis could have been performed for any window size or any other point on the CDF curve. This is emphasized here because in other types of applications it may be desirable to look at a specified level of the CDF function (e.g. CDF$\neq$ 0.5) or at a specified window size. Finally, the tests illustrated here hold over both the static and stress conditions.

Another question to be addressed is whether the effects of the physiological stress induce a performance change for each task number. Using the data from the seven subjects and four replications of the stress condition, Table V illustrates the effects of stress on tracking performance.

Table V Comparisons of Stress vs Static Conditions

Hypothesis test on μ values	$\bar{t}$ for this test	p
ff_1 stress > ff_1 static	3.34	<.05
ff_2 stress > ff_2 static	1.54	<.10
ff_3 stress > ff_3 static	2.81	<.05
ff_4 stress > ff_4 static	5.83	<.01
ff_5 stress > ff_5 static	3.14	<.05

The t statistic used in this test is the same as in equation 5. One can now see the impact on performance degradation as noted by the effect of stress on tracking in Table V.

In part II of this study, it was desired to study this sensitivity effect for the second class of tracking tasks. Table VI illustrates the actual error score results for Part II as well as the equivalent values found in Part I:

Table VI

	Part I				Part II			
	Static		Stress		Static		Stress	
ff#	$\bar{x}$ mean	$\bar{x}$ s.d.	$\bar{x}$ mean	$\bar{x}$ s.d.	$\bar{x}$ mean	$\bar{x}$ s.d.	$\bar{x}$ mean	$\bar{x}$ s.d.
1	94.8	26.4	127.7	29.6	22.67	3.87	37.73	7.14
2	162.4	48.7	179.0	33.2	38.42	6.71	59.14	12.02
3	285.9	49.4	306.3	29.7	54.32	11.40	79.91	14.88
4	388.6	22.9	414.7	37.4	106.32	26.63	121.77	34.89
5	624.5	23.9	649.5	26.4	223.09	38.82	239.96	48.58

The results of the statistical tests indicate a performance decrement under various conditions. The subjective data from the questionnaire are presented next.

Results From The Questionnaire

In the questionnaire, the subjects were asked to compare the relative difficulty of the task they were presently tracking with the previous task. Responses of "more difficult", "less difficult", "the same", or "couldn't tell" were then correlated with the task numbers presented. These results are displayed in table VII for Parts I and II of this study:

Table VII - Correlation of Task Numbers with Subjective Responses*

	Part I		Part II	
Subject #	# Correct/Total	% Correct	# Correct/Total	% Correct
1	24/25	96%	21/25	84%
2	23/25	92%	25/25	100%
3	24/25	96%	19/25	76%
4	25/25	100%	25/25	100%
5	23/25	92%	23/25	92%
6	23/25	92%	24/25	96%
7	23/25	92%	25/25	100%

The subjects also commmented that as they were presented tasks with higher forcing function numbers, the tracking tasks required more attention. This corresponds to the description of workload cited earlier in which higher performance requirements coupled with more stringent attention requirements increase workload.

SUMMARY AND CONCLUSIONS

This study used sensitive tracking tasks to evaluate performance degradation under acceleration stress. The tasks designed here had to satisfy certain criteria. First they had to be zero mean, constant variance, sum of sines. Second, open loop scores for all five tasks had to be identical. Third, the autopilot runs also had to yield a consistent score for all five tasks. When the human was tracking these tasks, however, a performance decrement had to be observed dependent on forcing function number for static tracking. In addition, the performance decrement had to occur as a function of the experimental conditions stress versus non-stress for each forcing function number.

At the conclusion of the experiment the subjects were given a questionnaire to rate the different tasks. Subjective ratings of each task in order of difficulty were necessary in order to verify the workload definition used here, which requires both a performance decrement and an attention requirement for arranging tracking tasks in order of increasing difficulty.

* Due to different subject pools in Parts I and II of this experiment, subject #N in Part I may not be the same person as Subject #N in Part II (N=1,.,7).

REFERENCES

1. Cooper, G. E.,1957, "Understanding and Interpreting Pilot Opinion", Aeronautical Engineering Review, March, pp. 47-56.

2. Cooper, G.E. and Harper, R.P., Jr., "The Use Of Pilot Rating In The Evaluation of Aircraft Handling Qualities" AGARD Report 567, April, 1969.

3. Repperger, D.W., Ward, S.L., Hartzell, E.J., Glass, B.C., and Summers, W.C., 1979,"An Algorithm To Ascertain Critical Regions of Human Tracking Ability", IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-9,No. 4,April, pp.183-196.

4. Jex, H.R., McDonnell, J.E., and Phatak, A.V., "A Critical Tracking Task For Man-Machine Related To The Operator Effective Delay Time I: Theory and Experiments With A First Order Divergent Controlled Element", NASA CR-616, October, 1966.

5. Verplank, W.L., 1977, "The Facilitating Effects of Uncertainty In Long-Term Manual Control", Proceedings of The Thirteenth Annual Conference on Manual Control, June 15-17, MIT, pp. 101-105.

6. Repperger, D.W., Rogers,D.B., "A Task Difficulty - G Stress Experiment", Submitted for Publication.

7. Levison, W.H., 1979, Comments on Workload, at the Fifteenth Annual Conference on Manual Control, Wright State University, Dayton, Ohio.

8. Snedecor, G.W. and Cockran, W.G., 1967, "Statistical Methods", The Iowa State University Press, 6th. Edition.

9. Guilford, J.P., 1956, "Fundamental Statistics in Psychology and Education", 3rd. Edition, McGraw-Hill.

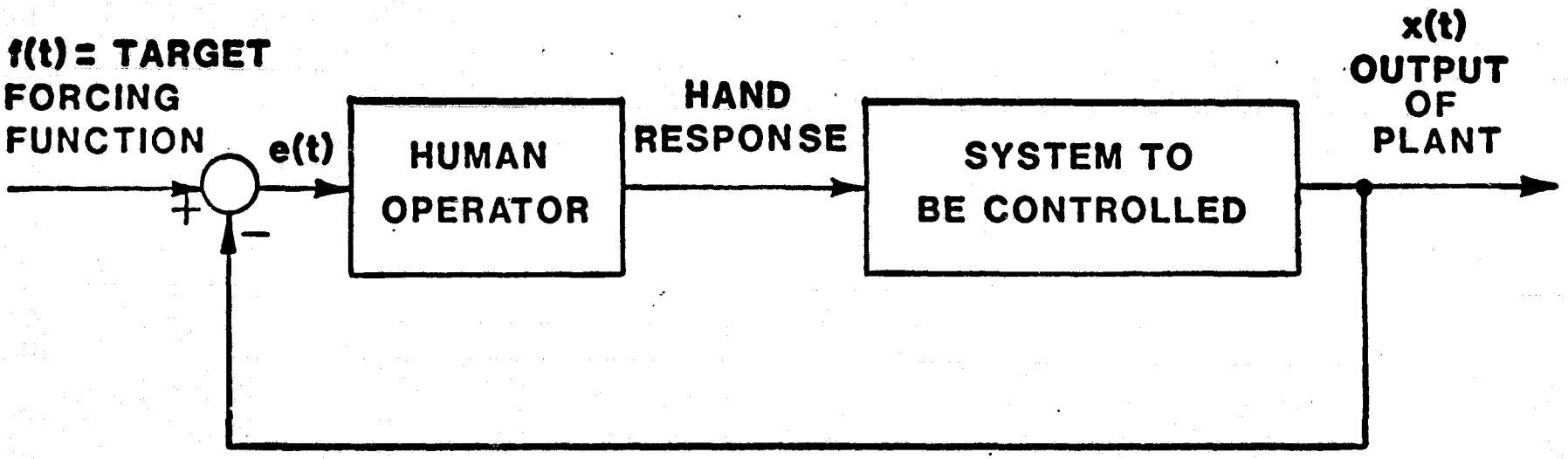


Figure (1) - THE MAN-MACHINE SYSTEM

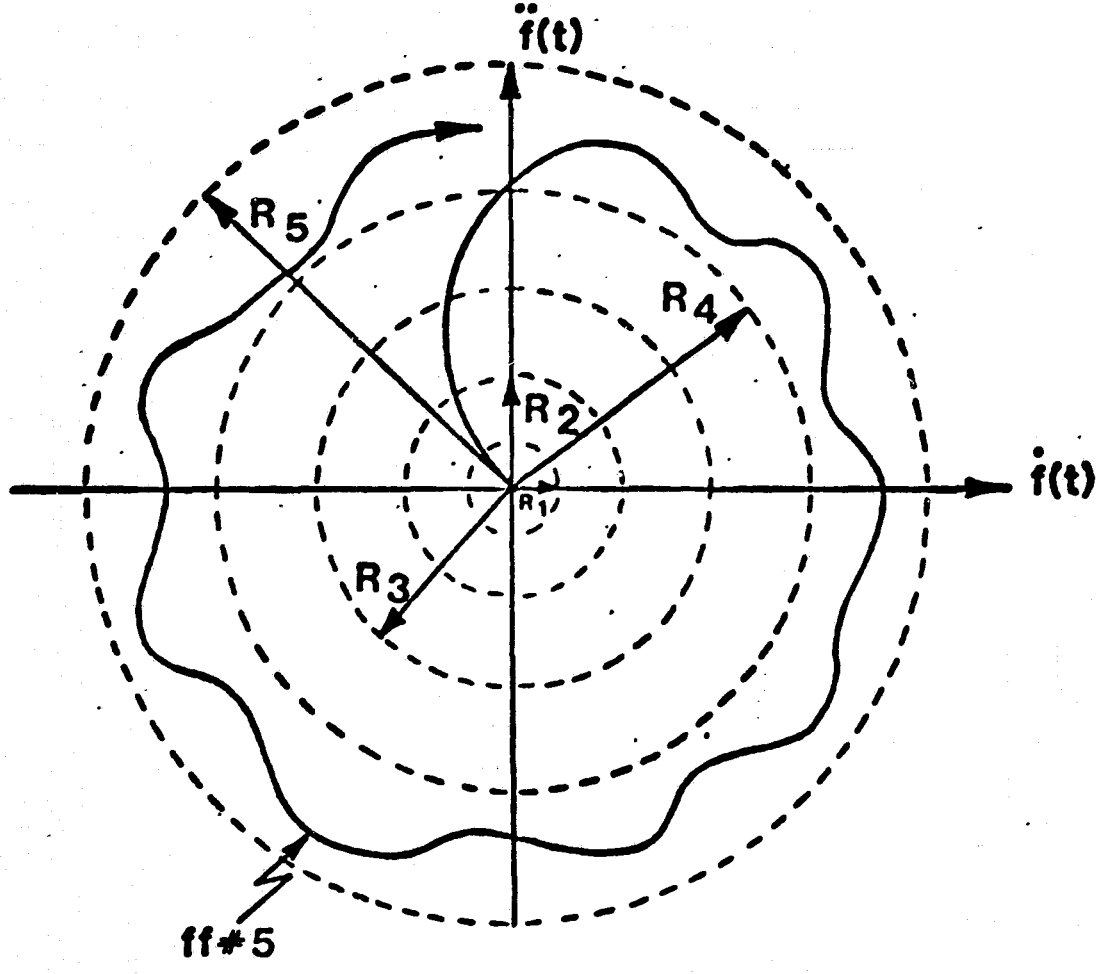


Figure (2) - A METRIC OF TRACKING DIFFICULTY

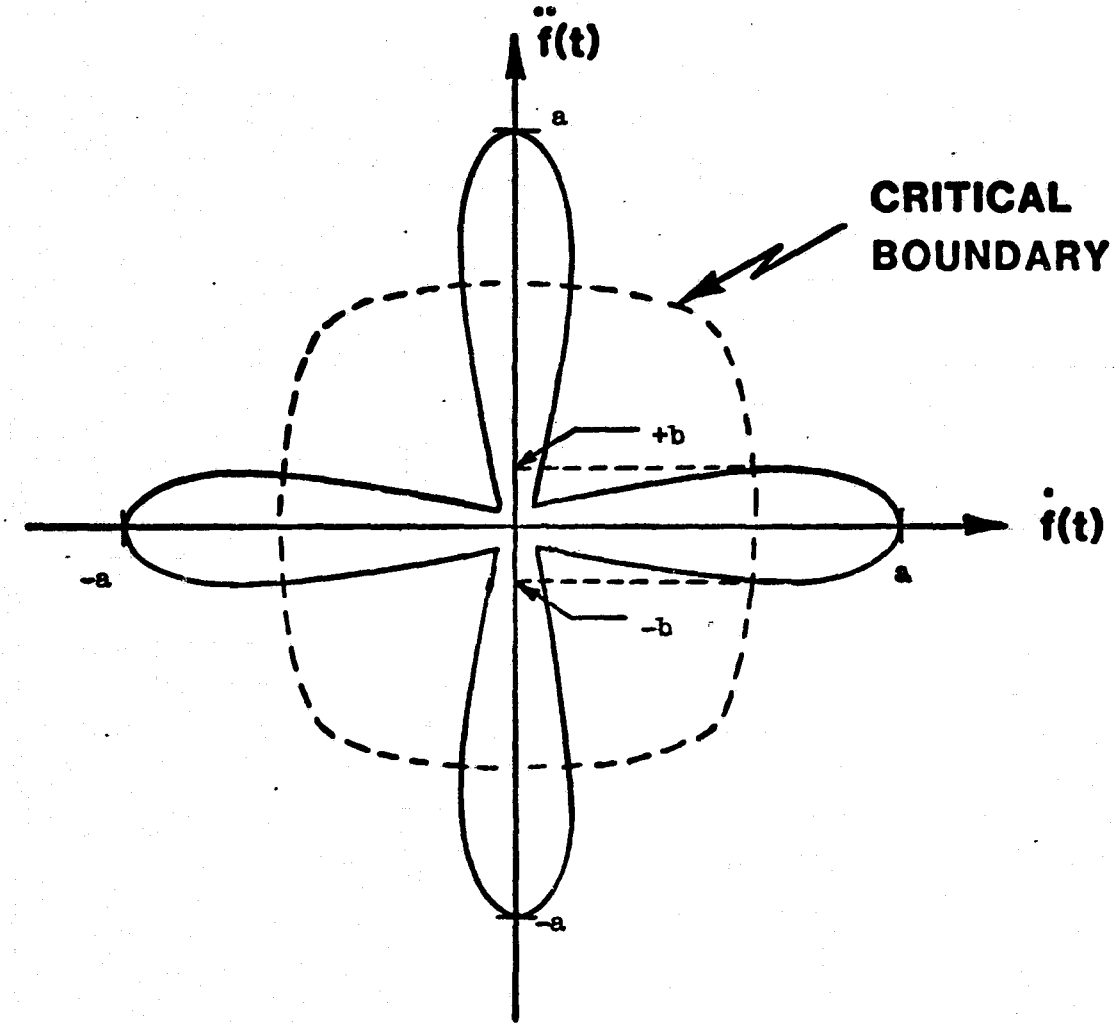


Figure (3) - SWEEP FUNCTIONS TO INVESTIGATE CRITICAL BOUNDARIES

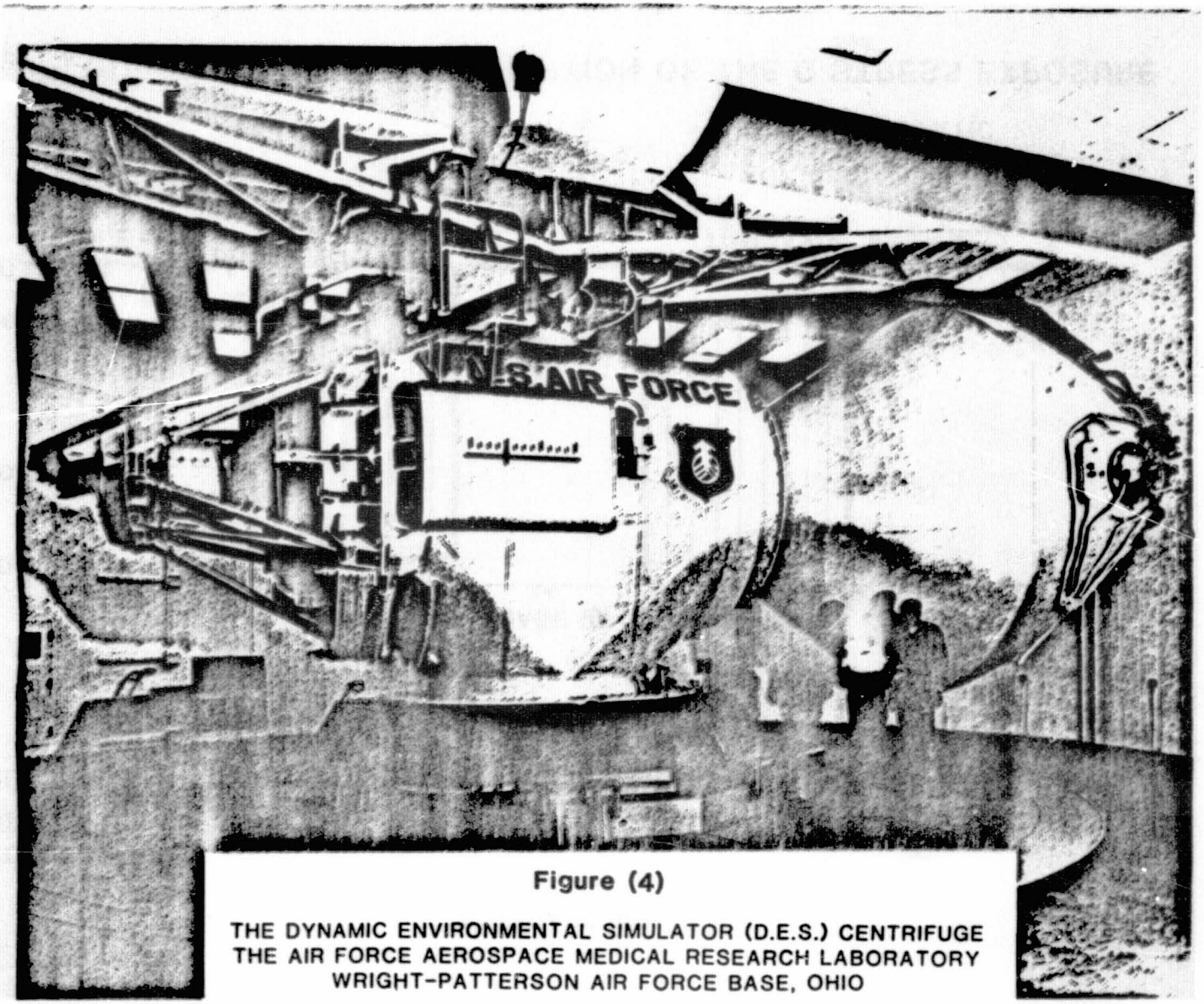


Figure (4)

THE DYNAMIC ENVIRONMENTAL SIMULATOR (D.E.S.) CENTRIFUGE
THE AIR FORCE AEROSPACE MEDICAL RESEARCH LABORATORY
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

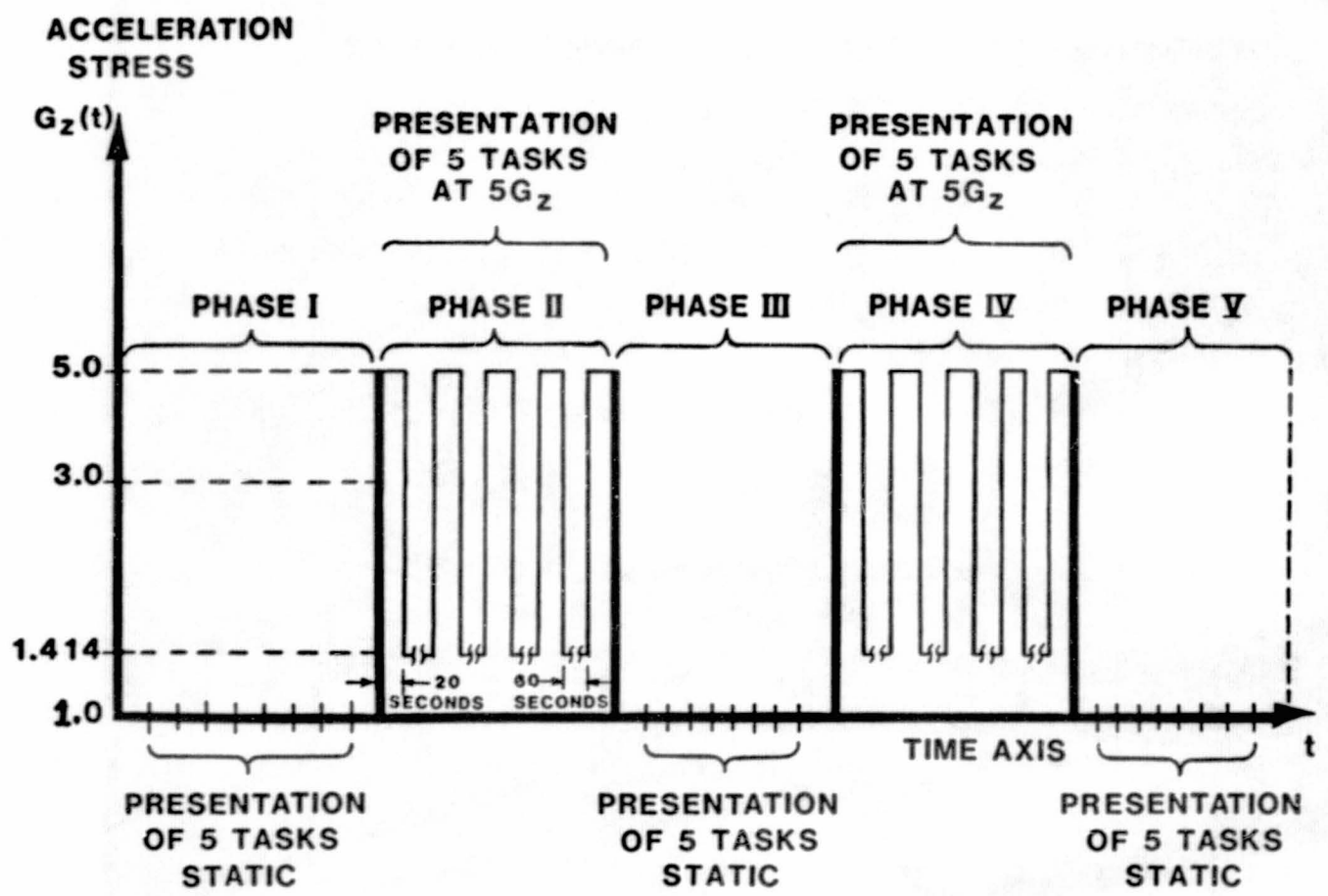


Figure (5) - A TEMPORAL DESCRIPTION OF THE G STRESS EXPOSURE

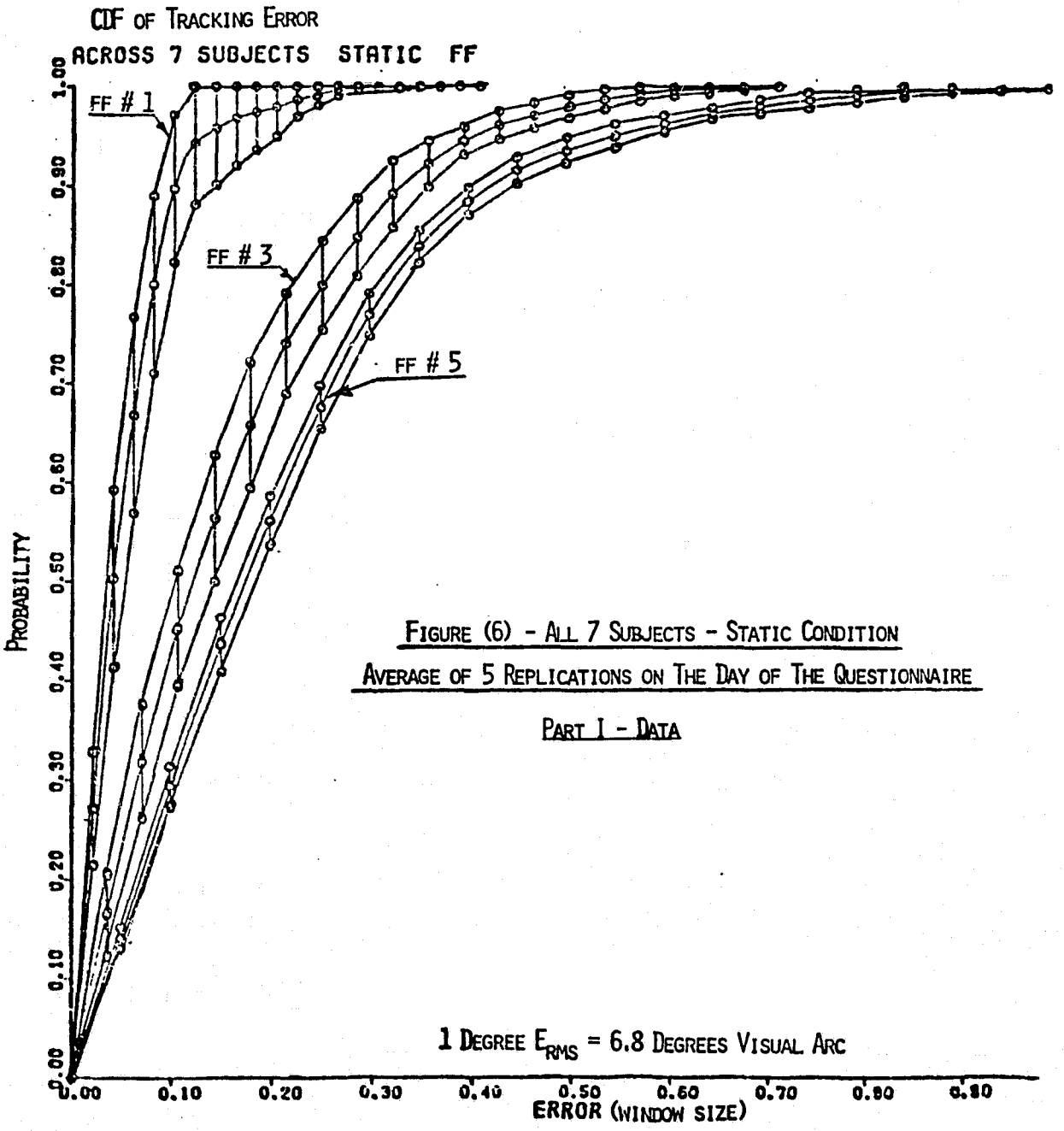


FIGURE (6) - ALL 7 SUBJECTS - STATIC CONDITION

AVERAGE OF 5 REPLICATIONS ON THE DAY OF THE QUESTIONNAIRE

PART I - DATA

SUBJECTIVE RATING SCALES AS A WORKLOAD ASSESSMENT TECHNIQUE

Kathleen L. Bird*

Virginia Polytechnic Institute

SUMMARY

Any investigation of the task workload inherent in flying must address many dimensions including cognitive, perceptual, and psychomotor. The present study employs a multidimensional bipolar-adjective rating scale as a subjective measure of operator workload in the performance of a one-axis tracking task. The rating scale addressed several dimensions of workload, including cognitive, physical, and perceptual task loading as well as fatigue and stress effects. Eight subjects performed a one-axis tracking task (with six levels of difficulty) and rated these tasks on several workload dimensions. Performance measures were tracking error RMS (root-mean square) and the standard deviation of control stick output. Significant relationships were observed between these performance measures and skill required, task complexity, attention level, task difficulty, task demands, and stress level.

INTRODUCTION

There is little agreement among scientists in how they conceive workload. To arrive at a functional, accurate, definition of workload several questions must be addressed. Does workload refer to the task demands imposed on the operator, or is workload the operator effort required to satisfy those task demands? What role, if any, does operator fatigue, physical and mental, as well as emotional stress play in the operator's assessment of workload? Is an individual's assessment of workload level really an assessment of a combination of all these factors? Most of the current workload definitions focus on a single facet of this multidimensional area. Jahns (1973) defines workload as "...the extent to which an operator is occupied by a task" (reference 1). Focusing on task performance measurements, Levison, Elkind, and Ward (1973) define workload as "...the fraction of the controller's capacity that is required for him to perform a given task to some specified or criterion level of performance" (reference 2). In an attempt to express the multidimensional aspects of workload, Tennstedt (1973) defines it as "...a summation of such processes as perception, evaluation, decision-making and actions taken to accomodate those needs generated by influences originating within or without the aircraft" (reference 3). While Tennstedt's workload definition addresses several workload dimensions, it falls short of addressing those questions previously posed.

* This research was conducted at NASA-Ames Research Center and was sponsored by NASA grant NAG-217 to Virginia Polytechnic Institute and State Univ.

The multidimensional aspect of workload is demonstrable in the flight duties of a pilot. A pilot's flight duties may encompass facets of cognitive (e.g., in-flight computations, fuel consumption management), perceptual (e.g., monitoring instruments, kinesthetic cues), and psychomotor (e.g., manual control of the yoke, rudder pedals) aspects. The pilot must also encounter, and cope with, the effects of fatigue and stress (mental and physical).

To investigate the multidimensional aspects of flying, workload assessment techniques (both behavioral and physiological) should address the cognitive, perceptual and psychomotor dimensions, as well as measure operator fatigue, and stress. Wierwille (1979) has suggested that a fruitful area of research would combine the best of physiological measures with behavioral measures in a multivariate analysis as a function of workload (reference 4). One or more workload assessment techniques need to be developed that can reliably measure the multiple dimensions of workload.

A subjective rating scale may offer a promising behavioral workload assessment technique. Hicks and Wierwille (1979) compared workload measurements obtained from rating scales with those obtained from primary task performance, secondary task performance, occlusion, and physiological measures (reference 5). Specifically, the rating scale proved to be a sensitive measure of operator workload in the performance of an automobile driving simulation task. Jenney, Older, and Cameron (1972) reported "...encouraging findings as to the usefulness and validity of subjective magnitude estimates" (reference 6). They recorded hourly subjective estimates of fatigue, tension, and task difficulty in assessing workload levels involved in performing an information processing task. Borg (1971) employed a simple rating scale and reported good agreement between perceived exertion, and difficulty, and physiological indicators of effort (stress) (reference 7).

The purpose of this study was to develop and validate a multidimensional rating scale to assess pilot workload. Several dimensions of workload were addressed, including cognitive, physical, and perceptual task loading as well as fatigue and stress effects.

Subjective Rating Scale

The multidimensional rating scale included 15 bipolar adjective pairs, one or more pairs addressing each of the several workload dimensions. These bipolar adjective pairs dichotomized: 1. skill required (no skill-much skill), 2. task complexity (simple - complex), 3. attention level (extremely low - extremely high), 4. monitoring (none - constant), 5. task difficulty (easy - difficult), 6. controlability (easy - difficult), 7. my performance (unsatisfactory - satisfactory), 8. instructions (clear - confusing), 9. task demands (undemanding - demanding), 10. energy level (lazy - energetic), 11. stress level (low stress - high stress), 12. activity level (idle - busy), 13. fatigue (tired - refreshed), 14. task stability (predictable - unpredictable), and 15. interest level (bored - interested).

The scales appeared one at a time on a CRT (cathode ray tube). The adjectives were positioned at opposing ends of a vertical line, with the descriptor (e.g., skill required, attention level) positioned below the scale. Subjects assigned a subjective rating (scale 1 to 100) to the tasks by positioning a cursor along the vertical line.

To validate the multidimensional rating scale, it is necessary to have subjects perform a battery of tasks which concentrate on different aspects of workload and examine whether the rating scale accurately measures these aspects. Future studies will employ a battery of six to eight primary tasks (similiar to the Civil Aeromedical Institutes Multiple Task Performance Battery, MTPB) which will include cognitive, perceptual, and psychomotor components. The primary tasks selected will closely approximate tasks demanded in flying.

The present study examines the psychomotor aspects of workload. Subjects performed a one-axis compensatory tracking task with six levels of difficulty. They rated the six tracking tasks for degree of workload using the multidimensional rating scale.

Tracking Task

The task was a one-axis compensatory tracking task with a K/S plant. A random number generator provided a rectangular distribution of frequencies (bandwidth of 1.0, 1.5, 2.0 rad/sec) filtered through a second-order filter to produce the forcing function. The filtered output produced the movement of the cursor.

Difficulty was manipulated by varying the standard deviation (SD of 32, 64) and bandwidth (1.0, 1.5, 2.0). The following tasks were presented: 1. task 1 (bw 1.0, SD 32), 2. task 2 (bw 1.5, SD 32), 3. task 3 (bw 2.0, SD 32), 4. task 4 (bw 1.0, SD 64), 5. task 5 (bw 1.5, SD 64), 6. task 6 (bw 2.0, SD 64). Performance measures were the tracking error RMS (root-mean square) and the standard deviation of the control stick output.

The tracking tasks were presented on a CRT. The six tracking tasks consisted of a vertical line (5.56 cm) which randomly moved in a lateral direction. Maximum displacement of the cursor was 12.70 cm. The subjects task was to keep this cursor centered between two stationary vertical lines (2.11 cm) by means of a control stick right and left.

METHOD

Subjects

Five males and three females (aged 18 to 42) served as paid volunteers. These subjects had been previously screened for tracking ability to gaurantee a minimum amount of psychomotor ability. A pilot study

yielded a criterion score which the subjects were required to achieve before selection. All subjects were right handed.

Apparatus

This study was conducted in a small, sound-attenuated experimental chamber. The subject was seated before a CRT. The control stick was located on the right arm of the chair. The throttle was located on the left arm. Data aquisition was recorded and task presentations were programmed through a Digital Equipment Corporation PDP-12 computer.

Procedure

Subjects were told the purpose of the study, given a description of the required tasks, and instructions for rating the tasks on the various workload parameters. They were told these tasks would vary in degree of difficulty. The importance of maintaining an equally high standard of performance across all tasks was stressed. To familiarize the subjects, tasks were presented (in order of ascending difficulty) and the subjects were permitted to track each task for one minute. During the experimental session the tasks were not presented in order of ascending difficulty but rather in random order. The subjects were given a one-minute practice session prior to each experimental session. The experimental session (for each tracking task) immediately followed the practice session for a duration of four minutes. After completing each tracking task subjects gave a rating for each of the 15 bipolar adjective pairs. As the scales appeared on the CRT subjects would move the throttle to position a cursor along the vertical line to indicate their rating. When they were satisfied with their rating, they pressed a response button. Immediately, a second scale would be displayed.

Subjects were required to perform each tracking task for a duration of four minutes. Standard deviation of the tracking error, output error RMS, standard deviation of control stick output, and stick output RMS were sampled every 30 milliseconds. The following analyses were performed on the data collected during the final two minutes of each experimental tracking session.

RESULTS

Effect of task difficulty on error RMS

A 2 (standard deviation) x 3 (bandwidth) analysis of variance was computed on the error RMS to determine if a significant difference in error RMS would appear as a result of manipulating the bandwidth and standard deviation. Results indicated a significant difference attributable to the bandwidth factor, F (2,14) = 49.30, $p < .01$ and the standard deviation

factor, F (1,7) = 311.59, $p < .01$. In addition, there was a significant interaction between the bandwidth and standard deviation factors, F (2,14) = 10.60, $p < .01$.

Effect of task difficulty on the standard deviation of stick output

A 2 (standard deviation) x 3 (bandwidth) analysis of variance was computed on the standard deviation of stick output to determine whether manipulating the bandwidth or standard deviation would produce a significant difference. There was a significant difference attributable to the bandwidth factor, F (2,14) = 8.25, $p < .01$ and the task standard deviation factor, F (1,7) = 59.81, $p < .01$. There was no significant interaction between these factors, F (2,14) = 2.24, $p > .05$.

Effect of task difficulty on bipolar adjective ratings

Fifteen 2 (standard deviation) x 3 (bandwidth) analyses of variance were computed on the ratings for each of the 15 bipolar adjective scales. There was a significant difference attributable to the bandwidth factor for the following scales: skill required (F (2,14) = 9.62, $p < .01$); monitoring (F (2,14) = 5.05, $p < .05$); task difficulty (F(2,14) = 12.59, $p < .01$); my performance (F (2,14) = 8.71, $p < .01$); task demands (F (2,14) = 4.46, $p < .05$); and stress level (F (2,14) = 5.90, $p < .05$). There was a significant difference in the ratings attributable to the standard deviation factor for the following scales: skill required (F(1,7) = 55.86, $p < .01$); task complexity (F(1,7) = 28.84, $p < .01$); task difficulty (F (1,7) = 18.04, $p < .01$); controlability (F(1,7) = 13.93, $p < .01$); my performance (F(1,7) = 7.67, $p < .01$); task demands (F(1,7) = 6.90, $p < .05$); stress level (F(1,7) = 7.94, $p < .05$); and fatigue (F(1,7) = 20.87, $p < .01$). No significant effect attributable to either the bandwidth or standard deviation factors were found for the following scales: attention level; instructions; energy level; activity level; task stability; or interest level ($p > .05$). There were also no significant interactions between the bandwidth and standard deviation factors for any of these analyses (with the majority of F values less than one, $p > .05$).

Relationship between error RMS and bipolar adjective ratings

To determine if a significant relationship exists between the error RMS and the scale ratings, Pearson product-moment correlations were computed (df = 47). The following significant correlations were derived between scale ratings and error RMS scores: skill required (r = +.55, $p < .01$); task complexity (r = +.40, $p < .01$); attention level (r = +.42, $p < .01$); monitoring (r = +.44, $p < .01$); task difficulty (r = +.57, $p < .01$); controlability (r = +.61, $p < .01$); task demands (r = +.51, $p < .01$); stress (r = +.28, $p < .05$); and task stability (r = +.32, $p < .05$). The remaining scales were not significantly correlated (at or above $p > .05$) with error RMS: my performance, instructions, energy level, activity level, fatigue,

and interest level.

Relationship between stick output standard deviation and scale ratings

To determine if a significant relationship exists between the stick output standard deviations and the scale ratings, Pearson product-moment correlations were computed (df = 47). The following bipolar adjective scales were found to be significantly correlated with stick output standard deviations: skill required ($r = +.55$, $p < .01$); task complexity ($r = +.53$, $p < .01$); attention level ($r = +.29$, $p < .05$); task difficulty ($r = +.55$, $p < .01$); my performance ($r = -.45$, $p < .01$); task demands ($r = +.46$, $p < .01$); stress ($r = +.40$, $p < .05$); and activity level ($r = +.29$, $p < .05$). The following scales were not significantly correlated with stick standard deviation: monitoring, controlability, instructions, energy level, fatigue, task stability, and interest level.

The correlation between error RMS and standard deviation of the stick output was significant ($r = +.44$, $p < .01$).

DISCUSSION

The relationship between increasing task demands and task performance (output error RMS) was examined. As the task demands increase (with an increase in input bandwidth and input standard deviation) subjects' ability to reduce this error decreased. Increasing the input standard deviations from 32 to 64 produces an increase in output error RMS. This result might be expected considering the relative amount of error the subject is asked to reduce. Increasing the bandwidth (1.0, 1.5, 2.0) produced an increase in output error RMS (Mean = 25.8, 36.2, 43.5). Manipulating the task bandwidth and standard deviation had a significant effect on the standard deviation of stick output. Doubling the input standard deviation (from 32 to 64) also doubled the mean standard deviation of the stick output (Mean = 12.3 (SD 32), Mean = 24.9 (SD 64)). A similiar increase in mean stick standard deviation could be attributed to an increase in the bandwidth (1.0 (mean = 15.7), 1.5 (mean = 19.7), 2.0 (mean = 20.6)). In summation, as the task demand increases, a degradation of task performance occurs. As task demand increases, subject effort (as measured by stick output standard deviation) also increases.

Theoretically, an increase in task difficulty should be reflected in the subject's evaluation of task workload level. Several rating scales are strongly related to output error RMS scores. Apparently performance degradation was strongly reflected in evaluation of skill required, task difficulty, controlability, and task demands. An increase in subject effort was strongly reflected in evaluation of skill required, task complexity, task difficulty, and task demands. These scales are most promising as indicators of certain workload dimensions and should be investigated further with other flight related tasks.

REFERENCES

1. Jahns, D.W. Operator workload: What is it and how should it be measured? In K.D. Cross and J.J. McGrath (Eds.) Crew System Design. Santa Barbara, California: Anacapa Sciences, July, 1973.

2. Levison, W.H., Elkind, J.I. and Ward, J.L. Studies of multivariable manual control systems: A model for task interference. NASA-CR-1746, 1971.

3. Wierwille, W.W. Pysiological measures of aircrew mental workload. Human Factors, 1979, 21, pp. 575-593.

4. Hicks, T.G. and Wierwille, W.W. Comparison of five mental workload assessment procedures in a moving-base driving simulator. Human Factors, 1979, 21(2), pp.129-143.

5. Jenney,L.L., Older, H.J., and Cameron, B.J. Measurement of operator workload in an information processing task. NASA-CR-2150, 1972.

6. Borg, G. Psychological and physiological studies of physical work. In Singleton, W.T., et al. (Ed.) Measurement of Man at Work. London: Taylor and Francis, 1971, pp. 121-128.

Physiological Assessment of Operator Workload During Manual Tracking : (1) Pupillary Responses

Qiuyan Jiang, Raja Parasuraman and Jackson Beatty

University of California, Los Angeles

SUMMARY

The feasibility of pupillometry as an indicator for assessing operator workload during manual tracking was studied. The mean and maximum pupillary responses of 12 subjects performing tracking tasks with three levels of difficulty (bandwidth of the forcing function were 0.15, 0.30 and 0.50 Hz respectively) were analysed. The results showed that pupillary dilation increased significantly as a function of the tracking difficulty which was reflected by the significant increase of tracking error (RMS). The present study supplies additional evidence that pupillary response is a sensitive and reliable index which may serve as an indicator for assessing operator workload in man-machine systems.

INTRODUCTION

The assessment of operator workload in complex man-machine systems has become more important for evaluating and enhancing the overall system performance as systems become more advanced and sophisticated. The traditional methods for evaluating the workload are subjective evaluation (Cooper, 1969), measures of primary and secondary task performance, and systems or engineering analysis. Results have shown that neither any single approach nor combination of methods can assess or predicate the workload adequately.

Physiological parameters, however, have manifested themselves as potential indicators in workload assessment (Wierwille and Williges, 1978). These indicators may reflect the dynamic change of the functional status of the operator induced by the task before or without performance decrement (Strasser, 1977). Of the physiological measures indicating momentary cognitive workload, pupillometric response appears to be most sensitive and reliable (Kahneman, Tursky, Shapiro, and Crider, 1969). For cognitive processing, multiple neurophysiological assessment has been emphasized recently since the hypotheses of multiple and hierachical resources were pointed out (Wickens,1980, Beatty, 1980). EEG seems to be a potential indicator both for general and specific brain activation if multiple channels of EEG could be analysed (Barth and Beatty, 1981, unpublished report). EEG studies during manual tracking show that the theta waves increase as a function of tracking difficulties (Jiang, Long, He and Zhao, 1977, unpublished report), lasting period and performance decrement (Kornfeld and Beat-

ty, 1977). It seems plausible to interpret this phenomenon as due to either an increase of tracking difficulty or as a result of trying best to attain the desired performance as the experiment demands. The study of psychophysiological mechanisms of theta waves indicates that they are related to information processing and mental effort (Schacter, 1977).

Task-evoked pupillary responses have been studied extensively in cognitive tasks and practical implications for operator workload measurement has been suggested (Beatty, 1980). As a part of the study of multiple assessment of operator workload, the effect of tracking difficulty on pupillary responses was studied in the present experiment and the feasibility of pupillometry for assessing operator workload was discussed.

METHOD

Twelve university students participated this experiment as subjects. The manual tracking system consisted of an auditory display, a finger joystick and a controlled element (K/S, simulated by analog computer).

The difficulty of the task was determined by the bandwidth of the forcing function. The cutoff frequencies were low (0.15 Hz), medium (0.30 Hz) and high (0.50 Hz) corresponding to easy, medium and hard tracking respectively. The period of time of tracking for each trial was 10 seconds starting with a warning signal, the word "ready". One second after this signal, the tracking forcing function (converted to a single tone) appeared in either of the two earphones. The pitch of the tone was proportional to the voltage of the forcing function or the tracking error. The subject was asked to control the stick laterally away from the ear in which the tone was heard. This in turn reduced the magnitude of the tracking error and the perceived pitch of the tone. Zero error corresponded to 500 Hz of the tone. The positive error was displayed in the left earphone and the negative in the right. For each level of difficulty of tracking, there were 20 trials of practice in order to attain a steady level of performance. After practice the data of pupillary response and tracking error (RMS) were acquired and stored in the computer. Because of a programming malfunction, the RMS tracking error data was obtained for only 9 of the 12 subjects. There were 20 trials of tracking for each level of difficulty. The order of these levels for the subjects to track was counterbalanced.

Vertical pupillary diameter was measured by a TV Pupillometer System for 10 seconds following the warning signal. The tracking error were recorded for 9 seconds during tracking. Pupillary responses were analysed using the program of this laboratory, and the mean and maximum amplitude of dilation during tracking relative to the 1 sec. baseline were computed.

RESULTS and DISCUSSION

Tracking performance: Tracking error (RMS) increased significantly with the bandwidth of the forcing function ($F(2,8)=17.85$, $P<.005$). This result is consistent with other studies indicating that the bandwidth of forcing function can be used as a factor to modify the tracking difficulty.

Pupillary responses: As shown in Fig. 1, the mean and maximum amplitude of the pupillary responses during tracking increased with tracking difficulty. These increases were significant ($F(2,11)=6.44$ $P<0.025$, for mean dilation; $F(2,11)= 6.13$, $P<0.025$, for maximum dilation). The mean pupillary dilations were 0.274 mm, 0.332 mm, and 0.377 mm, and the maximum dilations 0.377 mm, 0.440 mm, and 0.492 mm for the three levels of difficulty respectively. The peak dilations were approximately comparable to those evoked by short term memory tasks with 5, 6, and 7 items (Beatty, 1980), and the pupil dilation during easiest tracking even reached the magnitude evoked by a choice reaction time test with 8 alternatives (Beatty and Wagoner, 1980). The foundation of inter-task comparisons is based on the findings that the magnitude of pupillary responses during cognitive processing is independent of baseline pupillary diameter over a physiologically reasonable range of values (Bradshaw, 1969).

The grand averages of pupillary responses duing tracking over 12 subjects are plotted in Fig. 2, which shows the time history of the pupillary responses. It displays a stage of preparation for activation before tracking; a stage of fast dilation of pupil immediately following tracking; a stage of maintaining dilation; and a stage of a little constriction before stopping tracking. It is obvious that the effect of tracking effort on the pupillary response is dominantly reflected in the maximum difference but also in other stages. For instance the constricting stage lasted longer in the easy tracking just before rest and the magnitude of constriction was greatest in the hardest tracking case. Whether the latter phenomenon is due to the individual differences or due to overload of the task for some subjects as pointed out by Peavler (1974) should be tested by further study.

The available reports published so far on the physiological assessment of workload during tracking show that even though some physiological indicators are sensitive to significant differences between rest and tracking, they are unable to discriminate the different levels of tracking difficulty, for instance the heart rate variability (Hyndman and Gregory, 1975) and event related potentials (Isreal, Chesney, Wickens, and Donchin, 1980). The present result supplies additional evidence that pupillary response is a sensitive and reliable index for measuring mental effort.

As for the technique of the pupil recording, there is still a lot to be improved in order to be more convenient and less constraining for the subject, particullary for the operators performing tasks in a

real working condition. Owing to the great individual differences in the size of the eyes and the pupil, the color of the iris, and the ability to keep the eyes relatively fixated without blinking, the subjects should be selected appropriately.

This study only analysed the effect of auditory tracking on the pupillary responses, but the individual studies using visual display also revealed the tendency of pupillary dilation as the difficulty increased and residual attention lowered (McFeely, 1972).

One can conclude that pupillary responses can discriminate operator effort during tracking and can be compared with those evoked by other tasks. Furthermore, the innovation for the recording system is possible, hence, pupillometry may serve as a promising indicator to assess the operator workload in man-machine systems, particularly with additional indicators such as EEG or others which can reflect the activities of specific resources of cognitive processing.

REFERENCES

Beatty, J. A pupillometric index of operator workload. Technical Report No.23. UCLA.,1980.

Beatty, J., Wagoner, B. L. Pupillometric signs of brain activation vary with level of cognitive processing. Science, 1978, 199, 1216-1218.

Beatty, J., and Wagoner, B. L. Effects of response selection and task evoked pupillary dilations: Inplications for a motor theory of attention. Presented at Western Psychological Association, Hawaii, 1980.

Bradshaw, J. L. Background light intensity and the pupillary response in a reaction time task. Psychonomic Science, 1969, 14, 271-272.

Cooper, G. E. and Harper, R. P., Jr. The use of pilot rating in the evaluation of aircraft handling qualities. Technical Report NASA TN-D-5153, 1969.

Isreal, J. B., Chesney, G. L., Wickens, C. D., and Donchin, E. P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. Psychophysiology, 1980, 17, 259-273.

Hyndman, B.W and Gregory, T. R. Spectral Analysis of sinus arrythmia during mental loading. Ergonomics, 1975, 18, 255-270.

Kahneman, D., Tursky, B., Shapiro,D and Crider, A. Pupillary, heart rate and skin resistance changes during a mental task. Journal of Experimental Psychology, 1969, 79, 164-167.

Kornfeld, C. and Beatty, J. EEG spetra during a long term compensa-

tory tracking task. Bulletin of the Psychonomic Society, 1977, 10, 46-48.

McFeely,T.E. Pupil diameter and the cross-adaptive critical tracking task; A method of workload measurement. AD 749075, 1972.

Peavler, W.S. Pupil size, information overload, and performance difference. Psychophysilogy, 1974, 11, 559-566.

Schacter, D. L. EEG theta waves and psychological phenomena: a review and analysis. Biological Psychology, 1977, 5, 47-88.

Strasser, H. Physiological measures of workload - correlations between physiological parameters and operational performance. In K.E.Klein (Ed.), Methods to assess workload. AGARD Coference Print, No. 216,1977.

Wickens, C. D. The structure of processing resoureces. In R. Nickerson and R. Pew (Eds.). Attention and Performance VIII. New York: Erlbaum, 1980.

Wierwille, W. W. and Williges, R. C. Survey and analysis of operator workload assesment techniques. Technical Report S-78-101. Systemetrics Inc. Blacksburg, Va. 1978.

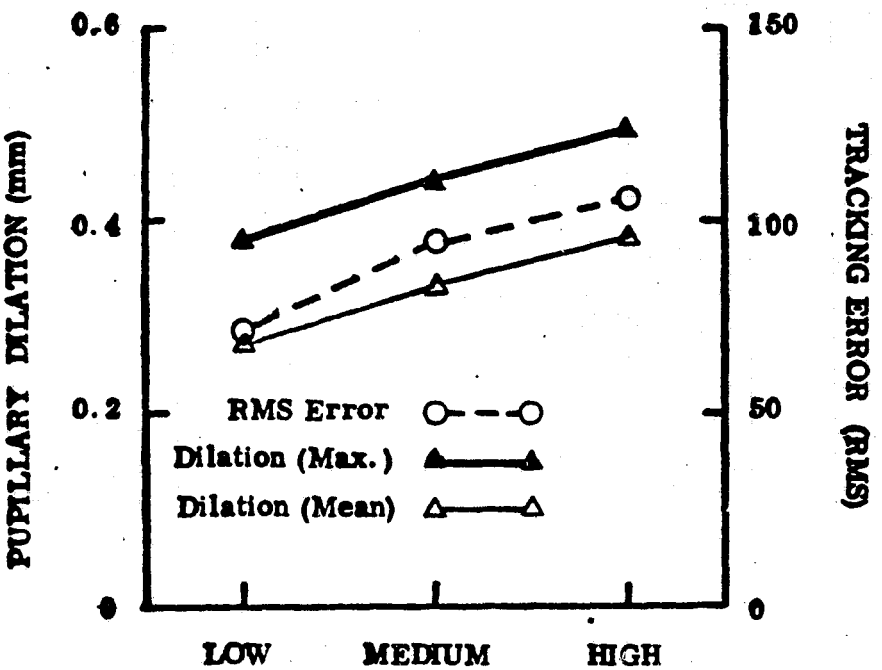


Fig. 1 Mean and maximum pupillary dilations and RMS tracking error (arbitrary units) for three levels of forcing function bandwidth (low, medium and high).

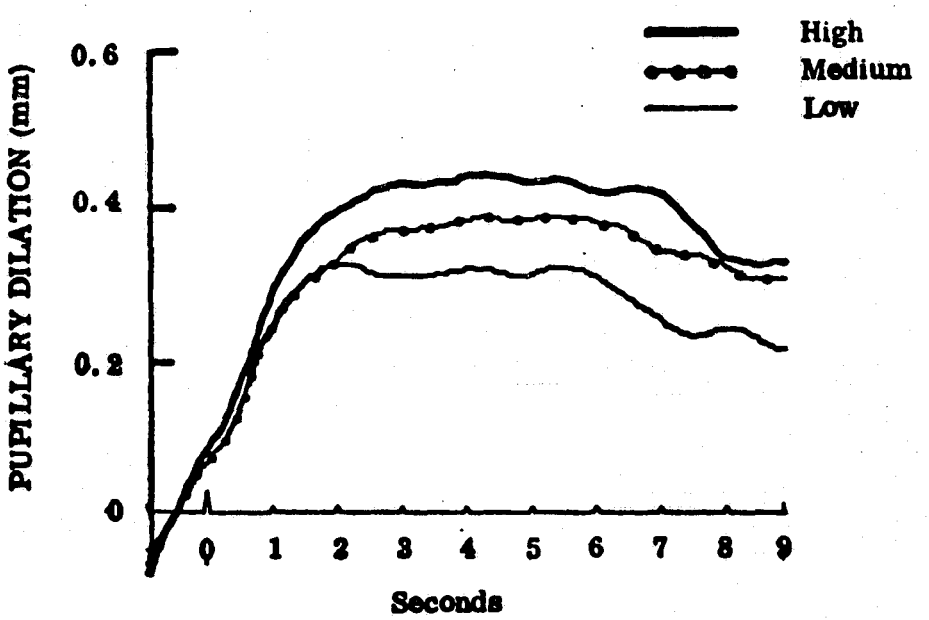


Fig. 2. Grand average of pupillary responses across 12 subjects for three levels of forcing function bandwidth (low, medium, and high).

omit

REGRESSION MODELS OF MENTAL ARITHMETIC

Renwick E. Curry

NASA Ames Research Center

ABSTRACT

Pilots of all categories of aircraft need to perform mental arithmetic calculations (time-to-station, ground speed, fuel at destination, etc.), and models of mental arithmetic are required for task analysis and performance assessment. An experiment was performed to test several regression models predicting solution time and errors as a function of the arithmetic operation and operands. Nine subjects solved 100 problems in each of two presentation conditions: the continuous condition showed the operands and operator as seen in normal paper and pencil solution; the discrete condition showed each operand separately and then the operation. There were significant solution time differences due to the type of arithmetic operation and presentation, and the variance of solution times was larger for the discrete presentation. The information measure of Thomas (1963), the total operation measure of Dansereau (1966), and individual operations (single digit add, carry, etc.) were used as independent variables to predict solution times and errors. The total number of operations and the individual operations were significantly better than the information measure in predicting solution times (average multiple R of .57, .59, and .40, respectively). Combined with the number of digits mentally stored, these measures provided d' values between error and nonerror trials of .74, .90, and .60.

REFERENCES

Thomas, H.B.G., Communication Theory and the Constellation Hypothesis of Calculation, Quart. Jour. of Exp. Psychol. 1963, 15, 173-191

Dansereau, D.F., and Gregg, L.W., An Information Processing Analysis of Mental Multiplication, Psychon. Sci., 1966, 6, 71-72

OMIT

EVALUATION OF SEPARATION OF PARAMETERS IN DISPLAYS IN APPROACH TO LANDING*

Garimella R. Sarma** and James J. Adams
NASA Langley Research Center
Hampton, VA

Fixed base simulation study was undertaken to ascertain the effects of separation or integration of essential display parameters in approach to landing. The display parameters used are bank, pitch, and heading angles as well as vertical and lateral displacement errors in a single engine high wing general aviation airplane in five degrees of freedom. Two military-type eight ball displays which have built-in crosspointer needles for displacement information were used for the purpose and the parameters are distributed in the two instruments by means of a switching box in a number of configurations. Two types of tests have been conducted in which eight pilots with varied amounts of experience participated. One test was with turbulence disturbance while the other was a step displacement correction.

Statistical analysis has been performed on turbulence test results with the configuration where all the parameters are displayed together in one instrument as the reference configuration. The results of these tests indicate that the lateral performance deteriorates significantly when the bank angle is separated from the rest of the parameters, and that it improves when bank angle and heading are displayed together. When the displacements are displayed together in a different location from the rest of the attitudes there is a significant improvement in performance over the reference configuration. No significant changes were observed in vertical control in any of the configurations.

Pilot aircraft system analysis indicates that the pilots reduce their bank angle and displacement loop gains in configuration where the bank angle is separated from the rest of the parameters, and restore these gains to almost that of the reference configuration when bank angle and heading are displayed together. The damping of the pilot-aircraft system became poor when the bank angle is separated and is the best when displacements are separated from the attitudes. It was also observed that a bank and heading angle combination gives rise to a better damped system in lateral control when compared to a pitch and bank angle combination.

---

*To be published by NASA

**NRC-NASA Research Associate

HORIZONTAL CONFLICT RESOLUTION MANEUVERS WITH A COCKPIT DISPLAY OF TRAFFIC INFORMATION

by Everett Palmer, Sharon Jago and Myrna DuBord

NASA Ames Research Center

PREFACE

This study investigated how pilots resolved potential conflicts in the horizontal plane when the only information available on the other aircraft was presented on a Cockpit Display of Traffic Information (CDTI). The intruder aircraft appeared on the CDTI with various programmed minimum miss distances, times to minimum miss distance, crossing angles, velocities, and turn rates. The pilot's task was to assess the situation and if necessary maneuver so as to avoid the other aircraft. No instructions were given on evasive strategy or on what was considered to be an acceptable minimum separation.

The results indicate that pilots had a strong bias of turning toward the intruder aircraft in order to pass behind it. In more than 50% of the encounters with a 90 degree crossing angle in which the intruder aircraft was programmed to pass either 2000 or 4000 feet behind ownship, the pilots maneuvered so as to pass behind the intruder. This bias was not as strong with the display which showed a prediction of the intruder's relative velocity. The average miss distance for all encounters was about 4500 feet.

INTRODUCTION

Two avionic systems are being developed which if implemented will provide pilots with information on other aircraft. One is the Cockpit Display of Traffic Information (CDTI) which shows the position and other information on nearby aircraft on an electronic map display. The second is the Collision Avoidance System (CAS) which in its simplest form provides warnings and maneuver commands to the pilot to avert possible midair collisions. In this latter system, to reduce the number of false alarms, commands are not issued until an immediate evasive maneuver is required to prevent physical contact between two aircraft. More complex collision avoidance systems also include an electronic display very similiar to a CDTI.

There are a number of potential advantages and disadvantages to presenting the pilot with situation information with a CDTI and command information with a CAS. The CDTI may interfere with the CAS by encouraging the pilot to assess the validity of alarms and commands and thereby exceed

the pilot response times that have been assumed in the design of the CAS algorithms. An assessment of the situation displayed on the CDTI may also result in the pilot making a maneuver different from that commanded by the CAS. On the other hand, the CDTI may aid the CAS if it allows pilots to detect potential problems before they become serious and make small maneuvers which will eliminate the collision threat. The CDTI may help reduce pilot response time to CAS commands by alerting the pilot to the need for a possible evasive maneuver (reference 1). The CDTI may also help eliminate the startle factor that might accompany a CAS alarm.

Flight research on collision avoidance systems without a CDTI has shown that often when pilots visually acquired an intruder aircraft by looking out the window before a CAS alarm sounded, they made an evasive maneuver different than that recommended by the CAS (reference 2). These differences usually arose when the CAS algorithm chose a maneuver that would maximize the minimum miss distance and the pilot chose a maneuver that would keep the other aircraft in sight and/or would allow ownship to pass behind the other aircraft. Situations where the pilot does not follow the command can cause problems if the other aircraft is receiving a complementary maneuver command. A CDTI display provides information in a very different format than that provided by the pilots' out-the-window view and these conflicts might not arise with this display. One obvious difference is that the intruder aircraft is always equally visible when displayed on the CDTI irrespective of distance or bearing.

The objective of this part task simulation study was to determine what types of maneuvers pilots made when the only information that they had on the other aircraft was that presented on a CDTI. Later studies will investigate pilot behavior when both CAS commands and CDTI situation information are simultaneously available.

METHOD

Display Hardware: The CDTI was displayed on a 18 cm (7") by 18 cm (7") cathode ray tube (CRT) located directly below the attitude indicator in a fixed base cockpit simulator. The center of the display was 0.44 rad (25 deg) below the horizontal and 90 cm (35") from the subject's eye reference point. The display elements were generated by a general purpose stroke writing computer graphics system.

Display Symbology: The ownship was always displayed with a curved ground referenced predictor. This predictor curves during a turn to show the effect of turn rate on future aircraft position. Figure 1 shows the two types of intruder predictors. They were: 1) the curved ground referenced (CGR) predictor which shows the future position over the ground with the provision that the current velocity and turn rate of the aircraft remain constant; (2) Curved ownship referenced (COR) predictor which shows the intruder's future position relative to ownship assuming both aircraft maintain constant turn rates and velocities. The following display elements were not changed throughout the experiment: 1) present position of ownship

was always indicated by a chevron symbol - the actual location of ownship was at the vertex of the symbol; 2) present position of the intruder was indicated by a dot in the center of a circular symbol; 3) RNAV route and runway symbols provided ground objects for background; 4) the width of the terrain displayed on the map was always 18.5 km (10 nm). With this map scale, 1.85 km (1 nm) on the ground equals 1.3 cm (0.5") on the display and 1.85 km on the map subtended a visual angle of 0.82 degrees; 5) the display was oriented with ownship track up; 6) track was updated every 0.1 seconds; 7) ownship position and all intruder information were updated every 4 seconds; and 8) ground referenced history which shows the past flight path of the aircraft over the ground was always present. No sensor noise or tracker lag was simulated.

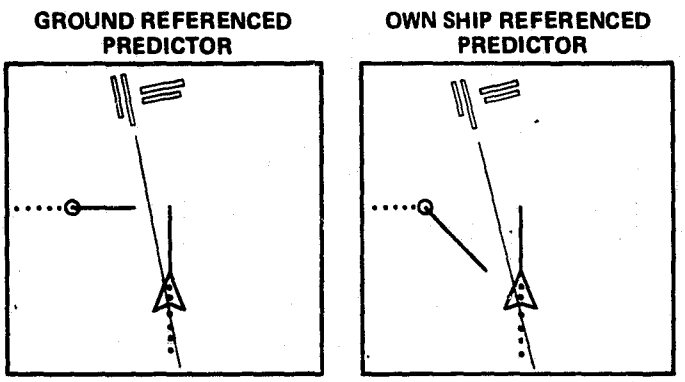


Figure 1. Two Experimental Display Formats

Encounter Variables: The experimental design incorporated a mixed factorial and star design. An encounter between the ownship and intruder ship was defined by the programmed miss distance (PMD), initial time to minimum miss distance (lead time), crossing angle of the intruder, speed of the intruder, and turn rate of both aircraft (see table 1). Ownship velocity was always 180 knots and both aircraft were always at the same altitude. An intruder with a positive programmed miss distance would pass ahead of ownship if ownship did not maneuver. The intruder was always initialized on the left side of ownship. With the intruder on the left, ownship actually had the right-of-way but the pilots were instructed that they should assume that the intruder had neither a CDTI nor a CAS and was not aware of ownships presence. On four of the encounters the intruder did maneuver 20 seconds after first appearing, but these turns were independent of ownships position.

Task: Subjects were asked to view the CDTI which depicted both ownship and the intruder. Subjects used the trim switch to turn ownship either right or left. Each "click" of the trim switch incremented the turn rate by 1.5 deg/sec.

Encounter No.	Programmed Miss Distance (feet)	Time to Programmed Miss Distance (sec)	Crossing Angle (deg)	Intruder Speed (knots)	Ownship Turn Rate (deg/s)	Intruder Turn Rate (deg/s)
1	+4000	30	90	180	0	0
2	+2000	30	90	180	0	0
3	0	30	90	180	0	0
4	-2000	30	90	180	0	0
5	-4000	30	90	180	0	0
6	+4000	60	90	180	0	0
7	+2000	60	90	180	0	0
8	0	60	90	180	0	0
9	-2000	60	90	180	0	0
10	-4000	60	90	180	0	0
11	+4000	90	90	180	0	0
12	+2000	90	90	180	0	0
13	0	90	90	180	0	0
14	-2000	90	90	180	0	0
15	-4000	90	90	180	0	0
16	+2000	60	90	120	0	0
17	-2000	60	90	120	0	0
18	+2000	60	90	240	0	0
19	-2000	60	90	240	0	0
20	+2000	60	45	180	0	0
21	-2000	60	45	180	0	0
22	+2000	60	135	180	0	0
23	-2000	60	135	180	0	0
24	+2000	60	90	180	+1.5	0
25	-2000	60	90	180	+1.5	0
26	+2000	60	90	180	-1.5	0
27	-2000	60	90	180	-1.5	0
28	+2000	60	90	180	0	+1.5
29	-2000	60	90	180	0	+1.5
30	+2000	60	90	180	0	-1.5
31	-2000	60	90	180	0	-1.5
32	+2000	60	90	180	0	+1.5*
33	-2000	60	90	180	0	+1.5*
34	+2000	60	90	180	0	-1.5*
35	-2000	60	90	180	0	-1.5*

* Intruder turned after 20 seconds

Table 1. Intruder parameters. In encounters 1 to 15, programmed miss instance and time to programmed miss distance were varied in a 3 by 5 factorial design. In the remaining encounters one parameter was varied at a time for PMD's of +2000 and -2000 feet.

Subjects: Eight current airline pilots served as paid subjects. They were selected from a pool of pilots who had volunteered to participate as test subjects.

Procedure: Pilots initially were given a brief description of basic CDTI concepts. A description of the predictors they would be using and instructions on how to use the trim switch to turn the ownship were given.

Pilots were instructed that they were to maneuver the ownship so that there were no collisions or near misses while keeping their deviation from course to a minimum. They were not instructed as to what was an acceptable miss distance or when or in which direction to maneuver. Pilots were instructed that the encounters would begin with different programmed minimum separations. It was explained that it was not mandatory to maneuver if it was thought that the current course was the best maneuver. The pilots were told that at the end of each encounter the CRT would display the minimum achieved separation. It was also explained that they would subjectively evaluate each encounter. Evaluation of encounters included ratings of how satisfied the pilot was with the overall maneuver.

Experimental Design: Pilots were presented with two blocks of all 35 encounters with one predictor type on the first day and one block of the same predictor on the second day. Four pilots saw each predictor condition in a between subject experimental design with subjects nested in predictor types. The order of the 35 encounters was random.

Objective data gathered included time and direction of the first turn, minimum achieved separation, maximum turn rate and maximum course deviation.

RESULTS

Analysis of the data from encounters 1 to 15, in which programmed miss distance and time to programmed miss distance varied while speed, crossing angle, and turn rate were held constant, was conducted by a four-way analysis of variance for repeated measures. Analysis of the data from the remaining encounters, in which the sign of the programmed miss distance, speed, crossing angle, and turn rate varied while the magnitude of the programmed miss distance and time to programmed miss distance was held constant, was conducted by a series of three-way analysis of variance for repeated measures.

Achieved Separation: Analysis of the data from all 35 encounters indicated that there was no significant difference in achieved separation between the two predictor types. Therefore, data from the two groups were combined.

The mean achieved separation over all programmed miss distances (PMD) and all times to minimum separation (lead times) in encounters 1 thru 15 was 4,410 ft. with a standard deviation of 1,560 ft. and a range of 300 ft. to 14,000 ft. This mean was fairly consistent regardless of the speed of the

intruder (encounters 16-19; 4,470 ft.), the crossing angle of the intruder (encounters 20-23; 4,700 ft.) or the turn rate of the intruder (encounters 28-31; 4,780 ft.). However, the mean achieved separation did increase by approximately 1,600 ft. when (a) ownship was turning (encounters 24-27), or (b) intruder initiated a turn after 20 seconds (encounters 32-35).

In encounters 1-15 the ANOVA found that the main effects of PMD and lead time were both significant. The highest mean achieved separation for each lead time occurred at the PMD of +4000 ft., while the lowest mean achieved separation for each lead time occurred at the PMD of -2000 ft. or 0 ft. (figure 2). The highest mean achieved separation for all PMD (except +2000 ft.) was at the lead time of 90 seconds while the lowest mean achieved miss distance for all PMD was at 30 seconds. It appeared that a greater difference in mean separation was found between the lead times of 30 and 60 seconds than existed between the lead times of 60 and 90 seconds. In addition, over all lead times, +PMD had a higher mean separation than -PMD. This result was found in all other encounter blocks except where the intruder turned after 20 seconds. In this instance the results were reversed.

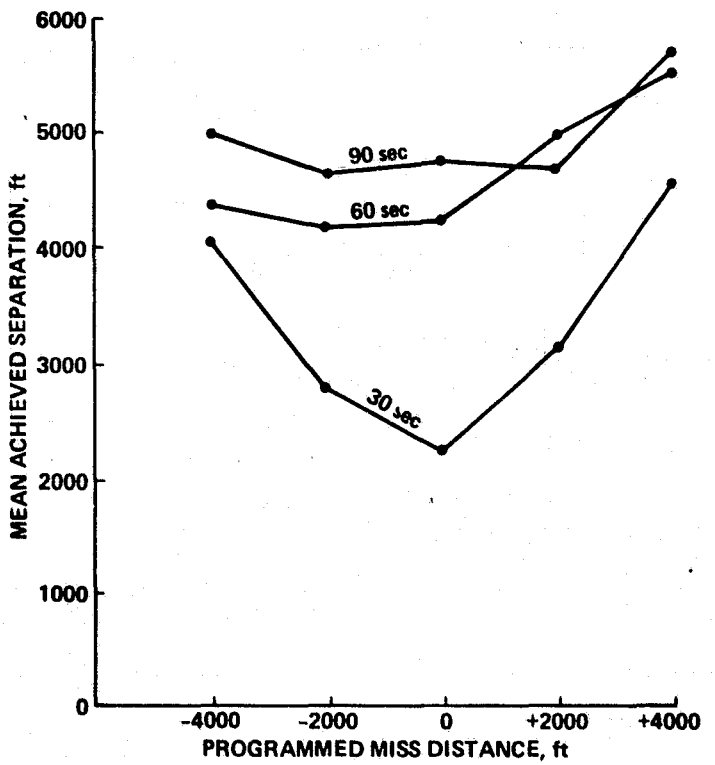


Figure 2. Mean Achieved Separation by Programmed Miss Distance for 30, 60, and 90 seconds to Programmed Miss Distance for Encounters 1 to 15.

Out of a total of 288 maneuvers in encounters 1-15, 31 maneuvers (10.7%) resulted in an achieved separation less than the programmed miss distance. These maneuvers will be referred to as "blunders". Over all lead times, the number of "blunders" was significantly higher for the -PMD (9%) than for the +PMD (1.7%) (figure 3). The number of blunders decreased as lead time increased but this effect was not significant. The CGR predictor group made significantly more "blunder" than the COR predictor group (figure 4).

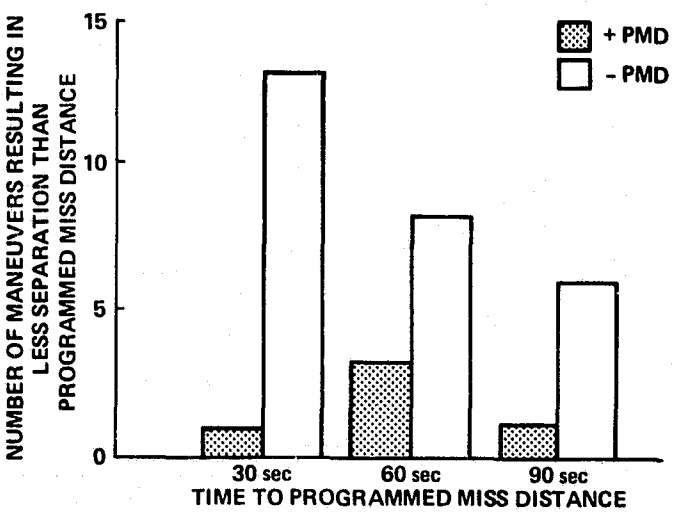


Figure 3. Number of Maneuvers Resulting in Less Separation than Programmed Miss Distance (blunders) by Position of Programmed Miss Distance (+PMD;-PMD) and Time to Programmed Miss Distance (lead time).

Latency Time Until First Maneuver: Data pertaining to this variable were obtained by recording the time from appearance of intruder on CDTI until pilots' initial maneuver. Trials in which no maneuvers were executed were not included in the analysis. Since analysis of data from all encounters except 16-19, in which intruder speed varied, showed no significant difference by predictor type, data for the two groups were combined in most instances. In encounters 1 to 15 the main effects of programmed miss distance and time to programmed miss distance were statistically significant ($p<.01$) and the interaction of programmed miss distance and time to programmed miss distance was significant ($p<.05$).

The latency time until first maneuver under the 30 second lead time condition (encounters 1-15) was 9.1 seconds with a standard deviation of 4.7 seconds. Thus, on the average, pilots made their first maneuver 9.1 seconds after the intruder appeared on the CDTI or after about 2 updates. The latency time until first maneuver increased as lead time increased.

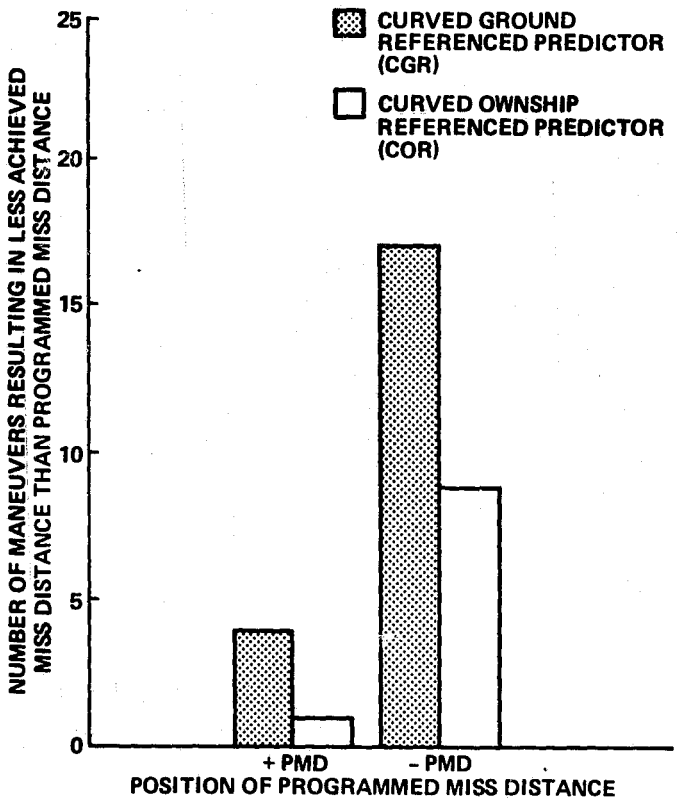


Figure 4. Number of Maneuvers Resulting in Less Separation than Programmed Miss Distance ("blunders") by Position of Programmed Miss Distance (+PMD; -PMD) and Predictor Type.

For each 30 second increase in lead time, pilots waited approximately an additional 10 seconds before maneuvering. This increase was fairly constant over all PMD.

In the encounters where lead time was held constant at 60 seconds but PMD, crossing angle, and initial turn rate were varied (encounters 7, 9, 20-31) the latency time until first maneuver ranged from 13.4 seconds to 19.9 seconds. However, when the speed of the intruder was varied (encounters 16-19) a significant difference ($p<.05$) was found by predictor type with the COR group, on the average, maneuvering approximately 10 seconds later than the CGR group.

Direction of First Turn: All pilots executed more left turns toward the intruder (67%) than right turns away from the intruder (17.5%) in encounters 1-15. Furthermore, seven out of eight pilots exhibited a statistically significant individual bias for turning left in the trials where there was a maneuver. This was consistent over all other encounters except where ownship

was already turning. In this instance, 58% of the responses were "no turn"; however, where a maneuver was executed it was in most cases in the same direction that ownship was already turning.

Data from encounters 1-15 revealed that the CGR predictor group made a higher proportion of left turns than the COR predictor group (figure 5). This trend was also evident in most other encounter phases.

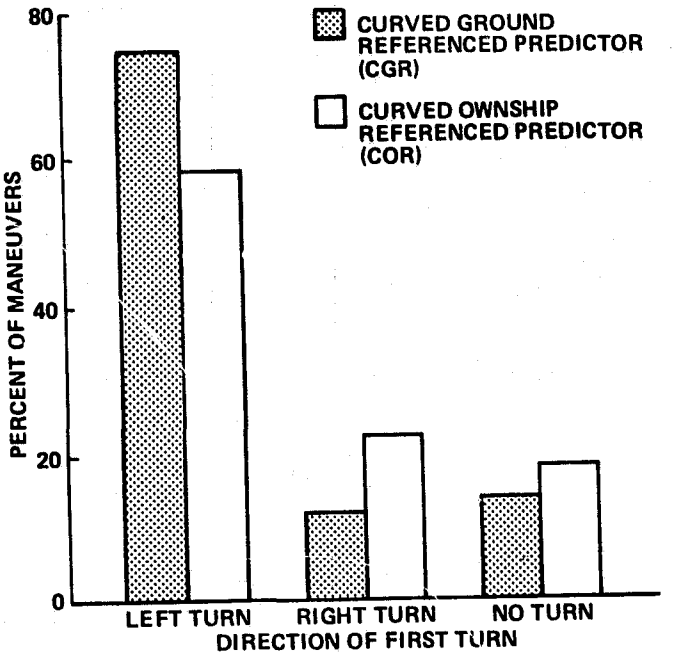


Figure 5. Percentage of Total Maneuvers which were "Left", "Right", or "No Maneuver" by Predictor Type.

Satisfaction Ratings: Analysis of data from the satisfaction ratings of each encounter (1=most satisfaction - 6=least satisfaction) revealed no significant difference in satisfaction by predictor type. Therefore, data from the two groups were combined. In encounters 1 to 15 the main effects of programmed miss distance and time to to programmed miss distance were significant ($p<.01$) as well as the interaction between programmed miss distance and time to programmed miss distance ($p<.01$).

The mean satisfaction rating for all lead time and PMD combinations for encounters 1-15 was 2.16 with a standard deviation of .52 and a range of 1.6 to 3.5. The pilots' lowest mean satisfaction occurred at the combination of 0 PMD and the lead time of 30 seconds. This was consistent with the point at which the lowest actual achieved miss distance occurred. The pilots' highest mean satisfaction occurred at the PMD of +4000 regardless of lead time.

Over all PMD, the pilots' mean satisfaction was lowest for the lead time of 30 seconds and about equal for the lead times of 60 and 90 seconds (figure 6). Once again, the mean satisfaction rating followed the pattern of the mean achieved separation.

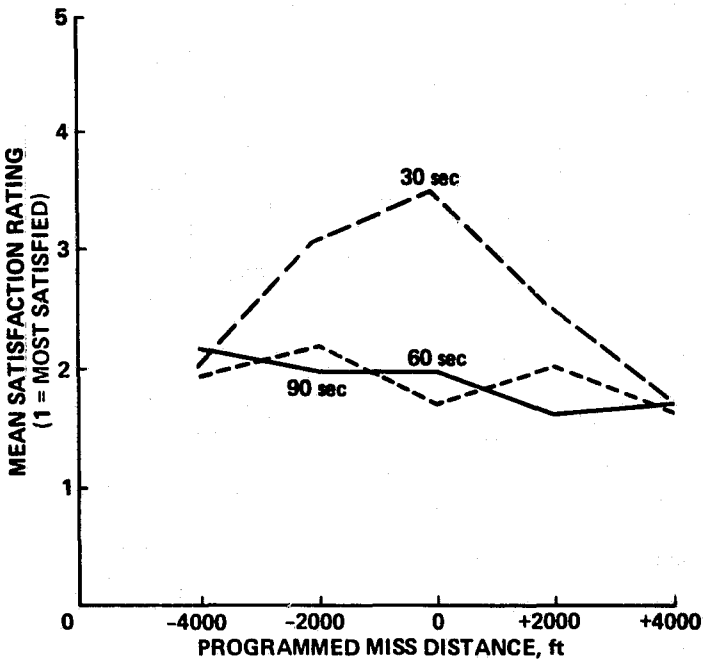


Figure 6. Mean Pilot Satisfaction Rating by Programmed Miss Distance and Time to Programmed Miss Distance (lead time).

This was evident for all encounter phases except where the intruder turned after 20 seconds. In this instance the results were reversed.

When the speed was varied (encounters 16-19), pilots were less satisfied at all speeds when the intruder passed behind ownship than when it passed in front of ownship. In addition, satisfaction decreased as speed increased. The lowest mean satisfaction occurred when the intruder was at 240 knots and passing behind ownship. No difference in satisfaction was associated with the crossing angle of intruder to ownship.

DISCUSSION

Seven of the eight pilots in this study had a bias toward turning so that they would pass behind the intruder aircraft. This bias existed even when the intruder aircraft was initially on a trajectory that would take it behind ownship if ownship did not change course. In these encounters, a

turn toward the intruder caused ownship to actually turn through a collision course with the intruder. A CAS algorithm that attempts to just maximize the minimum separation between two aircraft would always command a turn away from the intruder in these situations. Apparently these pilots' maneuver decisions were influenced by objectives other than just maximizing the minimum miss distance.

One objective expressed by the pilots was a desire to keep the other aircraft in sight. This was expressed even though no external vision was provided in this simulation and the CDTI display allowed the pilot to "see" equally well in all directions. Pilots were apparently attempting to keep the intruder in a position so that if they could look out the window, they would be able to see the other aircraft.

A second possible objective, though not volunteered by the pilots, was to minimize the amount of time it took to resolve the conflict. A turn toward the intruder allowed the pilot to more quickly resolve the problem and to return to the original course. Since speed and altitude maneuvers were not possible in this experiment, a horizontal maneuver that turned away from the intruder would sometimes place ownship on a course parallel to the intruder with the intruder effectively blocking ownship from returning to its original course.

A third consideration was that when a turn was made toward the intruder, it was perceptually easier to judge when the conflict was resolved. A typical maneuver was to turn toward the intruder until the intruder was directly ahead of ownship and then roll out of the turn. As soon as the intruder was directly ahead of ownship and not heading toward ownship, the pilot knew that the intruder was no longer a threat. During a turn away when an attempt is made to pass in front of ownship it is not clear that the situation has been resolved until ownship is directly in front of the intruder. During this maneuver both aircraft are going in the same direction and it can take a long time to reach this position. The experiment described in reference 3 showed that pilots made better judgements in predicting whether an intruder would pass in front or behind their ownship as prediction time decreased. A turn toward the intruder reduces this prediction time whereas a turn away increases it.

These considerations suggest that there may be rational reasons that the pilots turned toward the intruder aircraft even though this maneuver requires initially turning through a collision course. These results suggest that if pilots use a CDTI to assess conflict situations and the command is to turn away from the intruder that pilots may make maneuvers opposite to the command. One possiblity for avoiding this problem would be for the CAS algorithm to command vertical maneuvers in situations where a turn away command would normally be issued. A second possibility is through the CDTI design. In this study the group of pilots with the relative predictor made fewer maneuvers to go behind the intruder when the intruder was on a course that would take it behind ownship than the group of pilots that had the ground referenced predictor. A third possibility is training pilots not to initiate collision avoidance maneuvers based on CDTI situation information. Instructing pilots that the other aircraft is receiving a complementary

maneuver command would probably be persuasive. The results of this experiment might have been quite different if the pilot knew that the intruder had a CDTI or CAS and might therefore maneuver to avoid ownship. However if pilots feel that the CAS generates to many false alarms, they will probably start using CDTI and external visual cues to make independent threat assessments and evasive maneuvers in spite of training to the contrary.

CONCLUSIONS

Pilots were biased toward maneuvering so as to pass behind the intruder even when the intruder was programmed to pass behind their aircraft.

Of the two predictor display types, the pilots with the relative predictor display made fewer turns to pass behind the intruder, waited longer to make their first maneuver, and achieved about the same separation at closest approach.

Pilots maneuvered approximately 10, 20 or 30 seconds after the intruder appeared on the CDTI when the intruder was initialized 30, 60 or 90 seconds from the point of minimum miss distance.

Pilots generally expressed satisfaction with minimum achieved horizontal separations of 4000 to 5000 feet.

Future experiments would investigate the interaction between situation information displayed on a CDTI and maneuver commands from a collision avoidance system. Other experiments will investigate whether pilots can use the situation information on a CDTI to make small maneuvers well before a CAS alarm would sound that will resolve potential conflicts without triggering an alarm. These experiments will also investigate how pilot behavior changes when the intruder is a piloted simulator with a CDTI and/or collision avoidance system.

REFERENCES

1. Verstynen, H.A.: Potential Roles for the Cockpit Traffic Display in the Evolving ATC System. Society of Automotive Engineers, Inc. (SAE-800736), Warrendale, Pennslyvania, 1980.

2. Andrews, J.W.; Koegler, J.C.; Senne, K.D.: IPC Design Validation and Flight Testing (ATC-73). Final Report. prepared for the Federal Aviation Administration (FAA-RD-77-30). Lexington, Massachusetts: Lincoln Laboratory, Massachusetts Institute of Technology, 1977.

3. Palmer, E.A.; Jago, S.J.; Baty, D.L.; and O'Connor, S.L.: Perception of Horizontal Aircraft Separation on a Cockpit Display of Traffic Information. Human Factors, 1980, 22, 605-620.

SOME EFFECTS OF FIELD OF VIEW (FOV) AND TARGET SIZE ON LATERAL TRACKING AT HOVER

By
Harry T. Breul
Research Department
Grumman Aerospace Corporation
Bethpage, New York

SUMMARY

An exploratory flight-simulator experiment examined the gross effects of several factors potentially important to the design of a visual display system for aiding VTOL pilots in the difficult task of landing on a small sea-control ship. Field of view (FOV) and target size were the primary variables examined for a lateral tracking task in a full motion 5 degree-of-freedom (DOF) hover simulation, mechanized on Grumman's Research Hover Simulator (RHS). Both angular-rate-command and translational-velocity-command control systems were considered as well as two cockpit locations, at the aircraft cg and 15 ft forward of the aircraft cg. Sixteen experimental conditions were examined by two pilots in 105 tracking runs. The mean absolute value (MAV) of tracking error was used to measure tracking performance, and cross spectral transfer function analysis was performed to determine the pilot's ability to generate good open-loop transfer function characteristics as a function of the experimental variables.

In general, it was found that FOV and target size can have a large effect on the pilot's ability to generate open-loop gain, and on his tracking performance.

INTRODUCTION

A crucial element in the success of the sea-control ship concept is the all-weather operations ability of high performance VTOL craft assigned to relatively small ships at sea. The present study was suggested by the following flight scenario:

> It's nightime, there is limited visibility, gusting wind, and a VTOL craft is trying to land on a small pad at the aft end

of a destroyer that is being tossed about by a heavy sea. The darkness and overcast conditions mask any visual information from the pilot's periphery, and the ship's swaying superstructure looms before him to mask any other visual information about his position relative to the "ground". In addition, the cockpit is far in front of the cg, thus exposing the pilot to confounding lateral directional cues while he attempts to track the bounding ship through the narrow aperture of the head-up display (HUD).

The problem, of course, is to define the control and display requirements for safe, effective operation. But first, new data are required to determine things like how to replace the information normally acquired by the pilot in his peripheral FOV, whether he needs that kind of information, or even whether pursuit type information is necessary for tracking the ship's motions, and how a variety of vehicle design and control parameters interact with visual cueing requirements.

The work described here was an exploratory experimental look at the gross effects of two important visual cueing parameters, FOV and target size, and how they affect performance of the kind of lateral tracking required for a landing of VTOL craft on a small sea-control ship. The experiment also included variations in cockpit location and control mode. Two cockpit locations were simulated, at the cg and 15 feet forward of the cg, because many modern VTOL designs place the cockpit well forward. Seated there the pilot might easily confound lateral motion of the cockpit produced by lateral motions of the aircraft with lateral motion of the cockpit produced by aircraft yaw, particularly in conditions of deprived visual cueing. Two aircraft control modes were also simulated: angular-rate command and linear-velocity command. These were chosen to cover the range of candidate control schemes for the next generation of VTOL. The relatively simple-to-mechanize rate-command system requires highly developed piloting skills compared to the more complicated velocity-command system, which even a novice can fly reasonably well.

SYMBOLS

ϕ, θ, ψ	Roll, pitch, and yaw angles about aircraft x, y, and z body axes respectively, DEG.
u, v	Aircraft velocities along the x and y body axes respectively, fps
L_δ , M_δ	Lateral and longitudinal side-arm controller gains respectively, DEG/SEC^2/DEG
N_δ	Rudder-pedal gain, Deg/SEC^2/DEG

X_u, Y_v	Drag terms
$L_{\dot{\phi}}$, $M_{\dot{\theta}}$, $N_{\dot{\psi}}$	Angular damping terms, SEC^{-1}
g	Acceleration due to gravity
$\dot{Y}_{GUST}$	Inertial reference lateral velocity of aircraft due to atmospheric turbulance, fps
Y_{TAR}	Inertial reference lateral position of target, ft
ϵ	Inertial reference lateral displacement between the target and the simulated aircraft
δ_{LAT}, δ_{LONG}	Lateral and Longitudinal displacement of side-arm controller, DEG
δ_{RP}	Rudder-pedal deflection, DEG

SUBSCRIPT -

N	Indicates a sample data quantity, eg $F_N = F(N\Delta t)$

SIMULATION TECHNIQUE

The experiment was performed on Grumman's 6-DOF Research Hover Simulator (RHS) (Fig. 1). Conceptually, the RHS is a continuous-rotation yaw platform, supported by three independently controlled "jacks" mounted on a cart that is driven around the floor by a large "x-y plotter" type mechanism. The three jacks (only two are visible in Fig. 1) impart the pitch, roll, and heave motion to the yaw platform. They are traction-type linear actuators that move up and down a rotating shaft with a speed proportional to shaft rpm. They produce extremely smooth motion with a frequency response that is "flat" out to 4.5 Hz. The other DOF employ more conventional hydraulic drive systems and have frequency responses good to 2-3 Hz.

The hover equations of motion were developed with an eye to simplicity. The experiment was performed in the context of lateral tracking at hover, but the primary variables were visual cueing parameters, rather than subtle variations in dynamic behavior. Thus, no attempt was made to rigorously emulate any particular vehicle dynamics or any particular control system behavior. Also, some simplifying assumptions were made to ease the burden on the small analog computer available for this study. A vector-supported vehicle was assumed with rotational drag about each axis and translational drag along each axis. The only coupling was produced by the horizontal component of the thrust

vector resulting from roll and pitch excursions. Small angle approximations were used. Because of the limited throw of the Z-axis on the RHS (± 1 ft), an aircraft vertical motion was not included, but the cockpit moved vertically due to pitching motions when the simulated cockpit location was forward of the aircraft cg. ALl control forces were applied as couples. The resulting simplified hover equations of motion for the simulated VTOL vehicle with an angular-rate-command control system are as follows:

$$\dot{u} = g\theta - X_u U \quad \ldots\ldots\ldots\ldots\ldots 1$$

$$\dot{V} = g\phi - Y_v V \quad \ldots\ldots\ldots\ldots\ldots 2$$

$$\ddot{\phi} = L_\delta \, \delta_{LAT} - L_{\dot{\phi}}\dot{\phi} \quad \ldots\ldots\ldots\ldots\ldots 3$$

$$\ddot{\theta} = M_\delta \, \delta_{LONG} - M_{\dot{\theta}}\dot{\theta} \quad \ldots\ldots\ldots\ldots\ldots 4$$

$$\ddot{\psi} = N_\delta \, \delta_{RP} - N_{\dot{\psi}}\dot{\psi} \quad \ldots\ldots\ldots\ldots\ldots 5$$

where: $X_u = Y_v = 0.176 \text{ sec}^{-1}$

$L_{\dot{\phi}} = M_{\dot{\theta}} = 4.587 \text{ sec}^{-1}$

$N_{\dot{\psi}} = 10.00 \text{ sec}^{-1}$

$L_\delta = 5.64 \text{ deg/sec/deg}$

$M_\delta = 7.44 \text{ deg/sec/deg}$

$N_\delta = 1.2 \text{ deg/sec/deg}$

Figure 2 shows the basic computer flow diagram for the v and ϕ DOF. With the function switches open v and ϕ are described by Eq. (2) and (3) for the angular-rate-command control system. The resulting time constants for ϕ and v are 0.22 sec and 5.68 sec repsectively. Closing the switches to feedback ϕ and v creates the simple translation-velocity-command control system used in the study ($K_5 = -3.3°/\text{fps}$ and $K_6 = 4.4°/\text{fps}$). The θ , u flow diagram, is nearly identical, increased controller gain being the only difference (K_1 becomes 34.128). The specific values for the coefficients in Eqs (1) thru (5), and the K_5 and K_6 feedback gains, used to produce the translational-velocity-command system, were arrived at empirically. For this purpose we were fortunate to have another large and detailed engineering hover-control simulation being performed at Grumman during the time this study was being formulated. We relied heavily upon comparison with that simulation, both analytically and through a pilot serving both studies, to insure that the simplified equations produced representative dynamics with the two control systems considered.

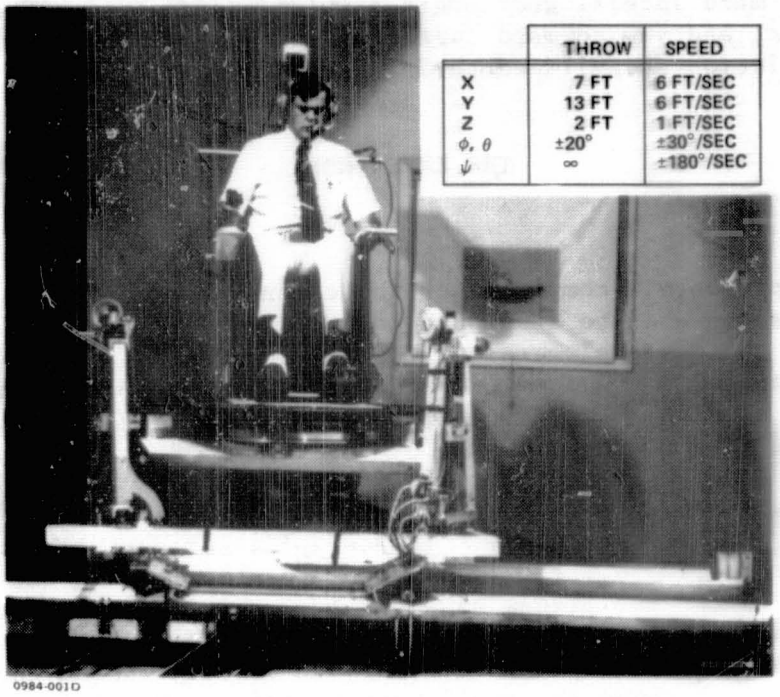


Fig. 1 Grumman's Research Hover Simulator (RHS)

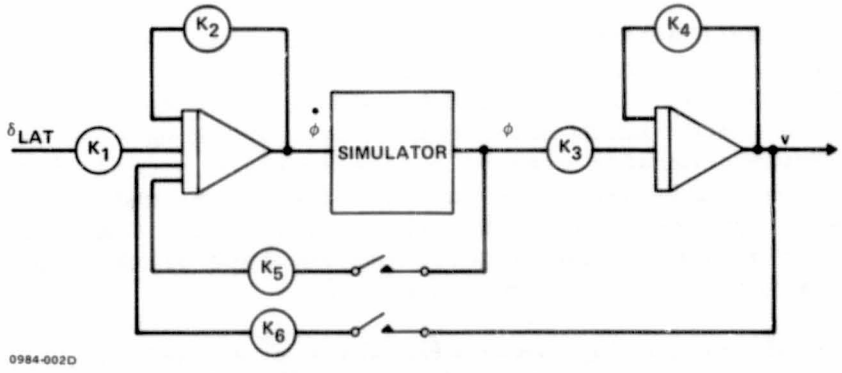


Fig. 2 Schematic Computer Flow Diagram for ϕ and v

The pilot made lateral and longitudinal commands with a 2-DOF side-arm controller, and yaw commands with rudder pedals. Both controllers had negligible friction and mild centering forces.

THE EXPERIMENT

Figure 3 presents a composite of the essential elements of the experimental setup used to study how lateral tracking at hover is influenced by two basic elements of the pilot's visual scene: FOV and target size. Two FOV conditions, "wide" (± 105°) and "narrow" (± 10°), were studied. They were achieved by adjusting openings in a cockpit hood fitted to the simulator. Two target sizes, "small" and "large", were used to provide two levels of background visibility. The small target was a black vertical rectangle (8 x 30 in) with a 1.0 in. wide white strip down the middle. It was split horizontally, with the bottom half projecting 8.0 in. in front of the top, giving the pilot the parallax between the two pieces as a cue for positioning his craft relative to the target. The background was a white wall with black vertical stripes (16 in. apart) standing immediately behind the target. The photo in Fig. 3 shows the small target against the striped wall. The large target was created by attaching a horizontal 4 x 8 ft sheet of tan foamboard to the rear of the small target. This masked the background wall and simulated the situation in which the pilot's FOV is dominated by the moving superstructure of a ship. Thus, tracking the large target through the narrow opening in the cockpit hood became a pure compensatory task (no background visible), while tracking the small target remained a pursuit task (background visible) for both FOV conditions.

The target moved from side to side in front of the cockpit with a motion like that of the port-to-starboard swaying of a landing platform on the stern of a small destroyer in a heavy sea. This kind of ship motion is characterized as having a lot of energy at a single frequency. Thus the targt drive signal was generated by adding the outputs of a sine-wave generator and a pseudorandom noise generator. Tracking runs lasted 204.8 sec and the noise generator created a line spectrum signal with a Δf of 1/204.8 Hz. The single sine wave was at 14/204.8 Hz, and the two signal generators were synchronized so that the target motion time-history was repeated identically every 204.8 sec. The envelope of the amplitude spectrum of target motion is shown in Fig. 4.

The detailed engineering hover simulation mentioned in the previous section was used to generate simplified gust disturbance data for use with the simplified vehicle simulation. The engineering simulation model was excited by a Dryden model (Ref 1) of atmospheric turbulence (RMS velocity = 4 fps) having a mean wind direction perpendicular to the nominal longitudinal plane. The resulting aircraft lateral inertial velocity was recorded and used as a gust-like disturbance in the present work by adding

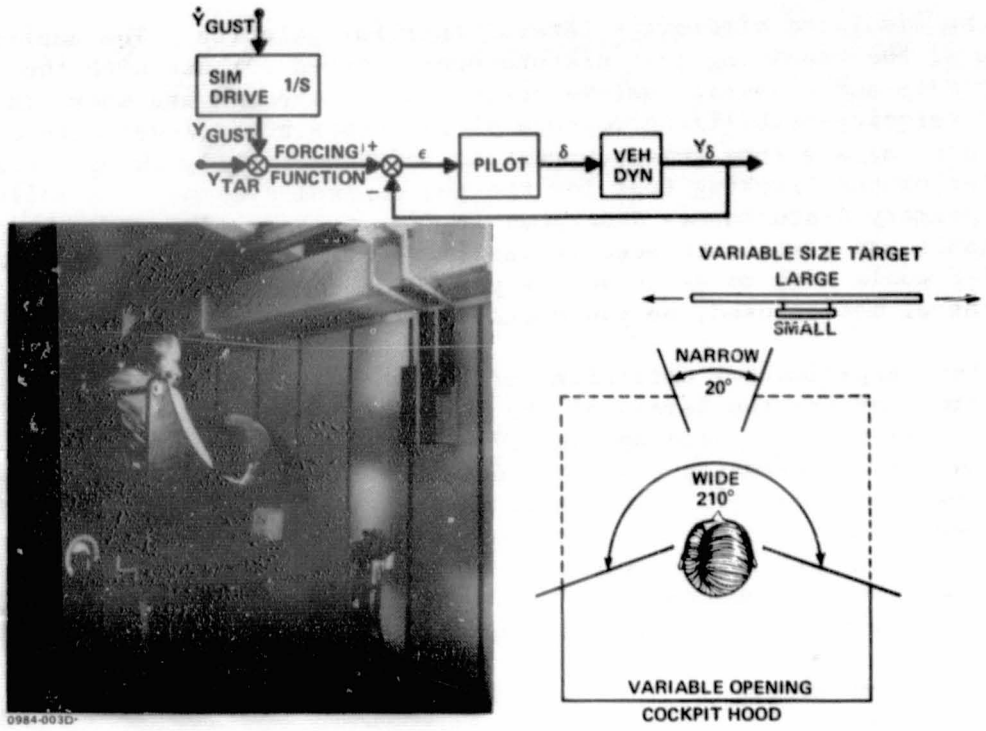


Fig. 3 Experimental Setup

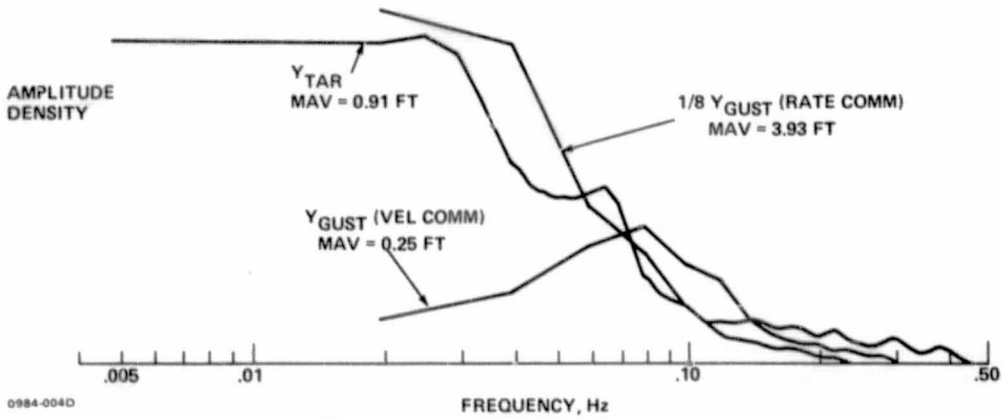


Fig. 4 Amplitude Spectra of Target and Gust Disturbances

it to the simulated aircraft's lateral inertial velocity. The amplitude spectra of the resulting gust disturbances created for use with the angular-rate and translational-velocity control systems are shown in Fig. 4. The velocity-stabilizing feature of the translational-velocity control system acts as a strong gust suppressant and dramatically changes the character of the tracking task for the two control systems. In addition to the primary disturbances described in Fig. 4, a small low frequency disturbance was also introduced in yaw (std dev = 7.1°/sec) to insure that the pilot would have to exercise the yaw DOF. With the very simple equations of motion used, no yaw motion would otherwise occur.

Sixteen experimental conditions are created by considering all combinations of the two levels of the four variables: FOV (wide and narrow), target size (large and small), control mode (angular-rate-command and translational-velocity-command), and cockpit location (at the cg and 15 ft forward of the cg). Two pilots made a total of 105 tracking runs at the 16 experimental conditions, and the order of presentation was randomized. One pilot was a recently retired (6 mos) Navy pilot with 800 hr fixed wing experience and 3700 hr in rotary wing craft, one-third of that gained operating off a ship at sea. The other "pilot" was a simulator engineer with over 20 years experience flying research and engineering simulations.

DATA AND ANALYSIS

During each 204.8 sec tracking run five variables were sampled at the rate of 10 samples/sec and stored on magnetic disc. They were: target position, Y_{tar}, tracking error, ϵ , lateral velocity due to gust, $\dot{Y}_{gust}$; lateral controller position, δ_{lat}; and yaw angle, ψ . The MAV and amplitude spectra of ϵ were calculated for all runs by both pilots. Pilot/vehicle open-loop transfer functions were also calculated, but for the helicopter pilot only. To do this, a sample data time history of aircraft lateral displacement due to pilot control, Y_δ (see Fig. 3), was needed. It was computed from the stored data as follows:

$$(Y_\delta)_N = (Y_{tar})_N + (Y_{gust})_N - \epsilon_N$$

where $(Y_{gust})_N$ was created by numerically integrating $(\dot{Y}_{gust})_N$. The process left a constant of integration unaccounted for, and the mean value of $(Y_\delta)_N$ in error. This was acceptable, however, because the harmonic analysis normalized the raw data by removing the mean.

DISCUSSION OF RESULTS

In this exploratory study, gross effects on performance were being examined, and the MAV of tracking error was the measure of performance used to compare the effects of the experimental variables. The averages of MAV tracking error for both pilots at each condition are summarized in Fig. 5. The "t" test (Ref 2), for measuring the confidence of differences between means, was applied to the data. MAV's significantly different (at the 5% level) are joined by brackets in the margin. Significant differences due to a single variable are indicated by brackets in the right hand margin and important significant differences due to more than one variable are indicated by brackets in the left hand margin. No comparisons were made between the angular-rate-command and the translational-velocity-command configurations. The shaded test conditions (4, 8, 12 and 16) all have the combination of narrow FOV and large target size. This results in completely masking all background information from the pilot's view, and changes the task from pursuit to compensatory tracking. These configurations produced an extreme degradation of visual cueing and are valuable as a sort of "benchmark", but, because they also produced a discrete change in task, they were not used for direct evaluation of the experimental variables. Therefore, significant variations involving the shaded configurations are not indicated in the figure.

CONTROL MODE	TEST CONDITION NO.	TEST CONDITION CONFIG*	MAV OF ϵ, IN.	NO. OF TRIALS	STD DEV
RATE COMMAND	1	WSA	6.9	4	0.42
	2	NSA	8.3	7	0.64
	3	WLA	7.7	9	0.51
	4	NLA	8.7	8	0.70
	5	WSF	7.8	7	0.64
	6	NSF	7.7	4	0.67
	7	WLF	7.9	6	0.67
	8	NLF	8.2	5	0.78
VELOCITY COMMAND	9	WSA	4.5	4	0.44
	10	NSA	4.7	10	0.24
	11	WLA	4.6	10	0.47
	12	NLA	4.2	3	0.38
	13	WSF	4.8	7	0.42
	14	NSF	4.7	7	0.31
	15	WLF	5.3	7	0.53
	16	NLF	4.6	7	0.30

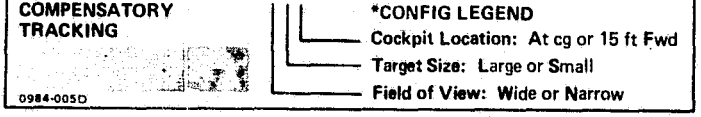


Fig. 5 Mean Absolute Value (MAV) of Tracking Error ϵ

For angular-rate-command control systems condition 1 is baseline (wide FOV, small target and cockpit at the cg). The best rate-command tracking performance is achieved at this condition. Comparison with condition 2 reveals a strong significant change in tracking performance due to reduced FOV. We hypothesize that it is a deterioration of inner-loop roll control due to a loss of roll-rate information from the pilot's peripheral FOV that leads to the poorer lateral tracking at condition 2. Roll and roll-rate information are still available in the pilot's foveal FOV because the striped wall is still visible behind the small target through the narrow opening. However, the loss of cues from the pilot's periphery is apparently crucial. Comparison of condition 1 with condition 3 shows that masking the background in the pilot's foveal FOV with the large target also produces a significant deterioration in tracking performance at condition 3. Here the pilot's peripheral information is not degraded and we suggest that neither are roll stabilization and control. Instead, we hypothesize that the lateral tracking suffers directly from the pilot's loss of information about the lateral motion in the outer-loop tracking task itself. That is, the edges of the large target are outside the pilot's foveal FOV and this reduces the precision with which he can visually sense pursuit-type information about target motion against the wall. Because tracking performance at condition 3 is also significantly different from tracking performance at condition 2 we can conclude that FOV has a stronger effect than target size.

The effect of moving the cockpit forward of the cg does not appear totally consistent. It clearly reduces tracking performance for the baseline configuration (condition 5 vs condition 1), but the combined effects of cockpit location and FOV (condition 6 vs condition 2) or cockpit location target size (condition 7 vs condition 3) are no greater than either effect alone.

In general, tracking error at the velocity-command conditions is much less than at the rate-command conditions, but it is not significant because the total forcing function (Fig. 3) was greatly reduced. This resulted from simulating the gust alleviation effect that is characteristic of velocity-stabilizing control systems by using a lateral inertial disturbance with much less energy (Fig. 4). What is of interest is the sensitivity of tracking performance with the velocity-command system to the experimental variables. For the translational-velocity-command control systems, condition 9 is the baseline (wide FOV, small target and cockpit at the cg). The first and obvious result is that neither reduced FOV nor increased target size significantly diminished tracking performance from the baseline (condition 9 vs conditions 10 and 11). If we believe, as suggested earlier, that FOV affects inner-loop roll stabilization, then we would not expect FOV to have an effect with the velocity-command control system, where the pilot is relieved of roll control. We would still expect target size to produce an effect on tracking performance, and suggest that the task has become so much easier (with velocity-command and the greatly reduced gust disturbance) that the effect is not critical (condition 9 vs condition 11). In this regard the effect of cockpit location is very interesting. The lateral tracking task

becomes more difficult when the cockpit is 15 ft ahead of the cg and the pilot must differentiate between lateral motion of the cockpit due to lateral motion of the aircraft and lateral motion of the cockpit due to yaw. This increased difficulty due to cockpit location does result in poorer tracking performance but only with the larget target (condition 11 vs condition 15). Also, this relatively poor tracking is improved when reverting to the small target with either FOV (condition 15 vs condition 14 or 13). Thus, we conclude that FOV does not affect tracking with the roll-stabilized-velocity command system and that target size does, at least at the more difficult forward location of the cockpit. This is consistent with the effects observed with the rate-command system and supports the notion that peripheral FOV information is needed for inner-loop roll stabilization and that foveal FOV information is needed for the outer-loop position tracking.

A limited amount of harmonic analysis was performed on some of the time-history data. Figure 6 is a plot of the amplitude spectral density of tracking error achieved at rate-command test conditions 1, 2 and 3 by the subject who is an experienced helicopter pilot. The curves are averages of from 3 to 6 repeats. The standard deviation shown is an aggregate for all curves. Test conditions 1, 2 and 3 demonstrate most dramatically the effects of FOV and target size. The plot shows a fairly uniform increase over the range of input frequencies (see Fig. 4) due to

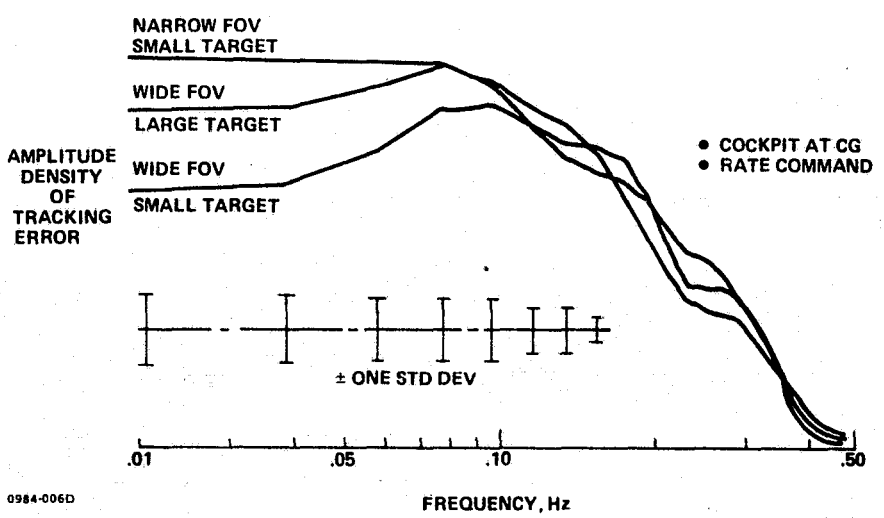


Fig. 6 Tracking Performance for Three Viewing Conditions

both decreased FOV and increased target size. Figure 7 is a plot of open-loop gain (see Fig. 3) for the same three conditions by the same subject. There was no variation in phase margin and the single plot shown is typical. The data indicate that the principle effect of both FOV and target size was to reduce the outer-loop position tracking, open-loop gain. We have indicated that the tracking error data suggest that FOV and target size affect different parts of the pilot's control activity. These curves show that the end result is to simply alter the open-loop gain.

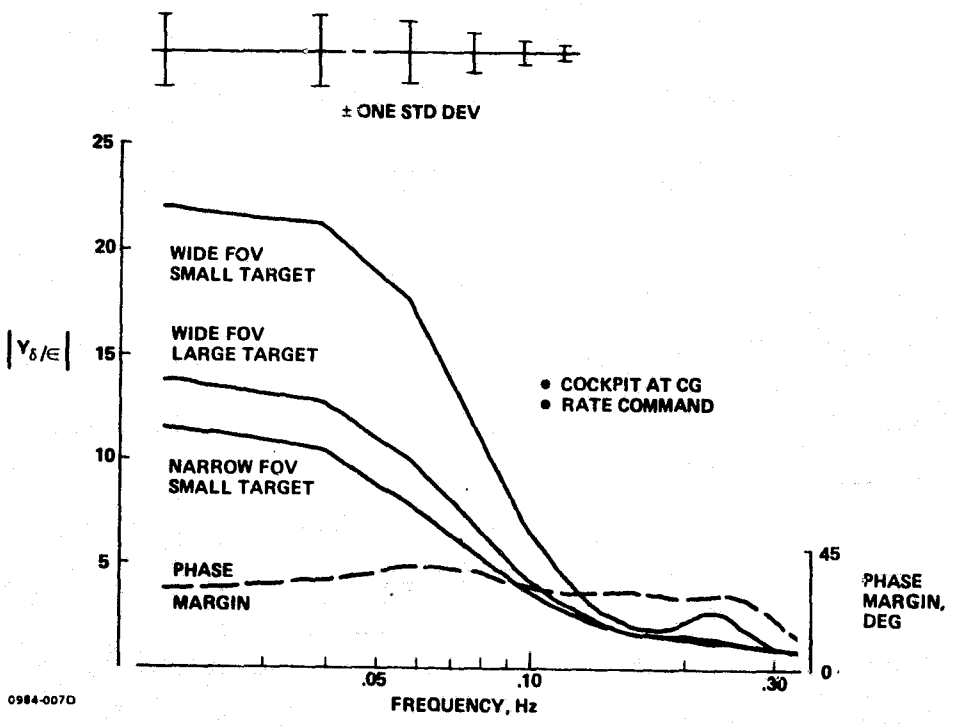


Fig. 7 Open-Loop Gain for Three Viewing Conditions

CONCLUDING REMARKS

An exploratory, full motion, 5-DOF simulator experiment using simplified dynamics has shown that FOV, target size, and cockpit location have a significant effect on a pilot's ability to perform lateral tracking at hover. Tracking error data support the hypothesis that the effects of FOV are largely separable from the effects of target size, FOV affecting inner-loop roll control, and target size affecting outer-loop tracking performance directly. Transfer function analysis suggests that both target size and FOV ultimately affect the outer-loop tracking performance by changing the outer-loop, open-loop gain the pilot can generate.

In a practical sense the results suggest that to achieve good lateral tracking performance at hover a pilot needs to sense roll information in his peripheral FOV or have the roll DOF stabilized. Normal roll information in the pilot's foveal FOV does not suffice. The results also show that added cockpit motion due to yawing about a cg aft of the cockpit can be detrimental, particularly when tracking large objects at close range, and suggests that the yaw DOF be stabilized for performance of analogous flight tasks.

REFERENCES

1. Chalk, C.R., et al: Background Information and User Guide for MIL-F-8785B (ASG), Military Specification - Flying Qualities of Piloted Airplanes. Air Force Flight Dynamic Laboratory Technical Report AFFDL-TR-69-72, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio, August 1969.

2. Hoel, Paul G.: Introduction to Mathematical Statistics, John Wiley & Sons. New York, 1947.

VIDEO FRAMERATE, RESOLUTION AND GRAYSCALE TRADEOFFS FOR UNDERSEA TELEMANIPULATOR CONTROL

Vivek Ranadive
Link-a-Bit Co.
San Diego, CA

Thomas B. Sheridan
MIT
Cambridge, MA

ABSTRACT

The high costs associated with a human diver working at ocean depths greater than 100m makes the use of remotely controlled visual inspection and manipulation attractive. However, long coaxial cables from the surface to the remote teleoperator cause additional difficulties. Therefore, sound communication without a tether is desireable. This form of telemetry, however, poses severe bandwidth restrictions so that its use for image transmission is in question.

The product of Frame Rate (F) in frames per second, Resolution (R) in total pixels and grayscale in bits (G) equals the transmission baud rate in bits per second. Thus for a fixed channel capacity there are tradeoffs between F, R and G in the actual sampling of the picture for a particular manual control task - in the present case remote undersea manipulation. A manipulator was used in the MASTER/SLAVE mode to study these tradeoffs. Images were systematically degraded from 28 frames per second, 128 x 128 pixels and 16 levels (4 bits) grayscale, with various FRG combinations constructed from a real-time digitized (charge-injection) video camera.[1]

When subjects first saw the video pictures with which they had to perform remote manipulation tasks, they refused to believe that they could succeed. Much to their surprise, however, they discovered that they were able to perform with a considerably degraded picture. It was found that frame rate, resolution and grayscale could be independently reduced without preventing the operator from accomplishing his/her task. Threshold points were found beyond which degradation would prevent any successful performance. It was observed that frame rate and grayscale could be degraded considerably more than resolution before teleoperation became impossible.

Isoperformance curves (curves of constant performance) were found for two subjects for various combinations of frame rate, resolution, and grayscale. These results were found to correlate closely with isotransmission curves (curves along which the information transmission rate is the same).

A general conclusion is that a well trained operator can perform familiar remote manipulator tasks with a considerably degraded picture, down to 50 K bits/sec, well below the several m bits/sec (5 MHz) normally used for broadcast video.

INTRODUCTION

Loss of life and costs approaching $5000 per working hour for deep ocean diving to do inspection and manipulation (as part of oil, gas, mineral and military operations) have motivated the development of remotely controlled vehicles having sensors and manipulators. Usually the operator observes through closed circuit video. (Sonic imaging is under development to provide a TV-like picture when water is too turbid for a video camera to work). Normally a high band-width coaxial cable is used as a communication link. However, such a cable can be extremely heavy for the submersible to drag around and can easily get caught in platform structure, rocks, etc.

Acoustic telemetry is an attractive alternative which avoids cable entanglement - at least for the lower portion of the signal path. Even though attenuation of sound in water is proportional to the fourth power of distance, several KHz for up to one Km distance are reasonable goals to send video pictures. The problem is: what kind of "real-time" image can be sent over a severely band-limited channel? More specifically, what are the best tradeoffs between frame rate, resolution and grayscale for an operator performing master-slave remote manipulation tasks?

SIMULATION OF UNDERSEA VIDEO IMAGING

Consider a remotely controlled submersible (teleoperator) with a video camera mounted on it, Figure 1. The bottleneck in this system is the information transmission capability from teleoperator to ship (or human controller).

A transmission channel of K bits/second could have

F frames per second

$R = \ell \times \ell$ pixels/frame resolution

G bits of grayscale/pixel = $\log_2$ (number of intensity levels) per pixel

Thus, the information transmission rate is

$$F \frac{\text{frames}}{\text{sec}} \times (\ell \times \ell) \frac{\text{pixels}}{\text{frame}} \times G \frac{\text{bits}}{\text{pixel}} = K \frac{\text{bits}}{\text{sec}}$$

Normal broadcast television has

$f = 30$ frames/sec

$\ell \times \ell \simeq 512 \times 512$ pixels/frame

assume $b = 4$ bits of gray

This means that normal broadcast television requires upwards of 30,000,000 bits/second. Current acoustic technology makes it possible to transmit from 30,000 bits/second in the shorter ranges to 3,000 bits/second for longer distances, up to 1 Km. It is, therefore, imperative for operators to learn to perform with coarse, slow pictures. Given the restricted bandwidth of operation, there must be compromises between the various features contributing to making up picture quality.

Common sense dictates that for some aspects of telemanipulation, framerate is most important while grayscale and resolution are not (e.g. moving a known object of good contrast while for other aspects high resolution with some grayscale is essential but framerate is not (e.g. identifying a fixed object). The literature provides little[2] on this, although a recent literature review by Cole and Kishimoto was helpful.

An experimental system called FRAG was developed to allow the experimenter to set the frame rate, resolution and grayscale of the video image. The system was a combination of commerically available components and special purpose (or home built) hardware.

The system used a General Electric TW2200 charge-injection camera. It contained not only the 16,384 pixels of the 128 x 128 camera, but all of the circuit logic necessary to perform a sequential raster scan and to generate synchronization signals. The CID arrays were fabricated as a silicon P.MOS device similar in many respects to some microprocessor and memory arrays.

The General Electric PN 2110 automation interface included analog or 8-bit digitized or thresholded binary video, power-clock signals, TTL signal level buffering and conversion and analog sweeps for CRT display presentation.

Variable framerate

FRAG was designed to continuously scan 28 frames/sec regardless of the selected sampling period since a slower speed integrated noise and saturated the picture. If the selected sampling speed was 28 frames/second then each frame could be displayed as it was sampled. For slower speeds, however, it was necessary to store frames in RAM and display each frame more than once. Thus, for a frame rate of 14 frames/sec, each frame would be displayed twice, and so on.

Variable Resolution

The GE interface had two registers of 128--x and y, corresponding to the horizontal and vertical scans. FRAG modified the counts from the x and y registers to result in the desired resolution.

Since the function of FRAG was to simulate as accurately as possible the effects of a low resolution picture transmission, it was decided to adopt a simple sampling scheme where every Nth pixel in both x and y was repeated N times. Thus, in Figure 2, a, b, c, and d would all be represented by A. This procedure required little real-time processing, and could be accomplished at fairly high speeds.

Grayscale

Four switches on FRAG produced 16, 8, 4, or 2 levels of gray by adjusting the coarseness of the digital-to-analog converters.

Example of tradeoffs

Using a cover picture of Life magazine, various combinations of resolution and grayscale are shown in Figure 3. For obvious reasons, frame rate cannot be shown in this manner.

Ranges of video adjustment

The FRAG system had the following ranges of operation:

Frame Rate: 0.109 to 28 frames/sec

Resolution: (8 x 8) to (128 x 128) pixels

Grayscale: 16 (at 2 levels) to 4 bits (16 levels)

Note that these were merely the available settings and were not necessarily usable by a human operator.

EXPERIMENT

Experimental Configuration

The teleoperator system including FRAG was set up as in Figure 4.

A modified Argonne E2 master-slave manipulator was used. It was attached to an Interdata computer which initialized the manipulator arms. This arm could move in all six degrees of freedom plus grasp. Although the manipulator was equipped with force-reflection, this feature was disabled so that the operator's only feedback came through the visual channel.

Experimental Tasks

The time required to complete a task under manual teleoperator control was expected to increase with the complexity of the task. The deterioration of the picture being used was expected to further increase the task completion time.

Tasks were selected to test manipulator performance based on the following criteria:

(i) task be representative of undersea manipulation. Such tasks include assessing damage, bolting/unbolting, connecting hoses, lifting objects, opening/closing valves, and reaching into confined spaces.

(ii) task performance be sensitive to required task accuracy.

In view of these criteria two tasks were designed for the teleoperator equipment

TAKE-OFF-NUT TASK (TON)

This task required the operator to locate a nut on a hub and then unscrew it. It was important not to drop the nut after removing it. The general method used by the subject was to grasp the nut, turn 180°, pull back to test if the nut was off, and if not release the grasp, reverse 180°, regrasp and repeat the operation. This task was representative of various "useful tasks" according to criteria (i).

OBJECT PLACEMENT TASK (123)

In accordance with criterion (ii) we chose a task which required fine positioning movements: to pick up a cork and place it sequentially in a preset sequence of three square areas on the table. Random 1-2-3 orderings were created so as to ensure the task being "closed loop", that is, to make visual feedback essential in moving between the three given squares in different trajectories.

Sound Feedback

Preliminary experiments showed that subjects made considerable use of sound feedback. For example, the sound of the manipulator colliding with the task hub became a valuable tool to determine position. In an ocean environment, such feedback probably would not be available. Accordingly, the lab airconditioner was turned on full to mask such feedback from the task.

Experimental subjects and their training

Two subjects were used as manipulator operators, both students in engineering. It was decided to compromise in the direction of well-trained subjects rather than use more subjects who were less well trained.

After gaining familiarity with the (force-reflecting) manipulator, force feedback was removed and the subjects were asked to do the TON task with direct vision. After some practice, the subjects were asked to perform the same task using a good (conventional) video picture. Finally, after the subjects were comfortable with this, the FRAG system and its accompanying degraded quality digitized picture was used instead of the conventional high quality video picture.

To insure that the subjects made definite progress during the training sessions, their performance was continuously monitored. The subjects had the best possible picture from the digitizer during this period. Figure 5 shows the learning curves for both subjects.

At the end of 10 intensive hours with FRAG, the two subjects' learning curves leveled in comparable fashion and the subjects were considered trained. The learning data were obtained on the basis of 10 trials for each data point value for each subject.

Experimental Protocol

The experiments were ordered so that two of the three variables (frame rate, resolution, and grayscale) were kept at a constant level while one third was varied.

As a performance baseline the best possible image conditions were used. All results were then compared with respect to this case. The best possible case had the conditions:

28 frames/sec frame rate

128 x 128 pixels resolution

4 bits grayscale

During the experiment, each subject was allowed to practice freely on each new image condition until "ready." Then twelve readings (i.e. time to accomplish tasks) on each were taken and the last six readings used as data. The TON task hub and 1-2-3 task paper were periodically reoriented to prevent the task from becoming rote.

RESULTS

Best Case

The average data for the best case was, surprisingly:

TASK	SUBJECT #1 TIME	SUBJECT #2 TIME
Take-off-Nut (TON)	28 seconds	28 seconds
Squares (1-2-3)	28 seconds	28 seconds

Relative performance measures P were defined as follows:

where T_n = TON time and T_a = 1-2-3 time

best case data = 28 seconds for both tasks

{all values averaged over six measurements}

$$P_n = \frac{28}{T_n} \times 100 \text{ for performance on TON task}$$

$$P_a = \frac{28}{T_a} \times 100 \text{ for performance on 123 task}$$

$$P = \frac{P_n + P_a}{2} \text{ for overall performance.}$$

Variable Framerate Results

The grayscale and resolution were kept constant and the framerate was varied. Figure 6 provides the results. This experiment was performed with two sets of grayscale/resolution settings: 1) 128 x 128 pixels, 4 bits gray

2) 64 x 4 pixels, 2 bits gray.

From these results it was clear that framerates below 5.6 frames/second considerably degraded performance and increased variability. At low framerates subjects had to use a "move-and-wait" strategy. Even though the move-and-wait strategy was time consuming, it worked!

Variable Resolution Results

The resolution was varied while maintaining a constant framerate and grayscale. Figure 7 provides performance curves for the variable resolution case.

It was found that the subjects successfully accomplished the nut removal (TON) task at resolutions as low as 32 x 32 pixels. It was noticed that the 64 x 128 pixel case resulted in considerably better performances than did 64 x 64. This showed that total number of pixels was more important than symmetry.*

*Symmetry was not entirely unimportant since 64 x 64 was definitely preferable to 128 x 32!

At lower resolutions, adequate operator training was important. A well trained operator could perform the task (after the manipulator is positioned) with little visual feedback.

Variable Grayscale Results

Keeping the framerates and resolution constant, the grayscale was varied. Results are shown in Figure 8. Notice that with maximum resolution (128 x 128 pixels) and maximum frame rate (28 f/sec), grayscale can be reduced to the two bit level without affecting performance. This is not true at a lower frame rate (14 f/s) where lowering grayscale does degrade performance.

DISCUSSION AND FURTHER ANALYSIS OF RESULTS

From the data gathered for these two specific tasks, it is apparent that the three parameters F, R and G could each be degraded, keeping the other two constant, without much effect on performance, up to a certain point where performance then degraded rapidly. Under the specified conditions: F = 28 f/s, R = 128 x 128, G = 4 bits, reducing the frame rate by a factor of 4 (2 bits) affected performance by only 20%. Similarly, reducing the grayscale alone by a factor of 2 bits reduced performance by 25%. In the case of the resolution, however, two bit reduction degraded performance of the TON task by 70%, while making the 123 task impossible to accomplish. It is especially useful to consider these facts in terms of the number of bits per second to be transmitted:

FRAME RATE*	RESOLUTION	GRAYSCALE	PERFORMANCE**	# OF BITS/SEC
28 f/s	128x128	4	100%	1,835,000
7 f/s	128x128	4	80%	458,750
28 f/s	128x128	1	75%	458,750
28 f/s	32x32	4	*33%	458,750

It is clear from these data that the number of bits per second could be kept the same and yet produce different performance for different combinations of F, R & G.

Correlation of Performance and Display Bit Rate

For the purpose of studying further the various trade-offs, "isoperformance curves" were constructed for combinations of F, R & G along which the performance was (almost) the same. Isotransmission curves (of constant information transmission rate) were superposed. These curves are shown in Figures 9, 10, and 11 respectively for RG with constant F, GF with constant R and RF with constant G.

The comparison shows, remarkably, that telemanipulation performance correlates very closely with bits per second of the display!

* Data combined for two subjects

** TON Performance, P not defined

Noise Problems at Low Framerates

It was noticed that at low framerates there was more noise in the picture. As a result of this, the operator's task was additionally complicated with a slower picture. From examination of the FRAG system, it was clear that the noise problems were not because of the charge-injection camera. It was therefore unclear whether:

(i) the noise was actually present in the slow images, or only

(ii) there appeared to be more noise in the slow pictures

The eye is often thought of as a low pass filter. This is in fact the reason why pictures (television, movies...) are shown not faster than 30 frames/second.

The fastest sampling rate of the FRAG was 28 f/s. Each of these frames was made up of signal and noise. In each consecutive frame the signal was the same* but the noise changed. At the higher frequencies the visual nervous system averaged the noise. At lower frame rates, however, each frame was displayed several times; here the frequency of the noise was now much lower. For example, at 28 f/s the noise had a frequency of 28 hz. At 1 f/s, the frequency was 1 Hz. Owing to the low-pass nature of the eye, the signal to noise ratio increased with frame rate.

Mathematical Model

Assume that the human eye averages over a period of τ seconds.

Let n = number of frames/τ seconds

F_n = nth frame in a particular τ period

S_n = signal in nth frame, all 1's white

W_n = noise in nth frame, all 0's black

$F_n = S_n + W_n(\sigma_n^2)$, where W_n^2 = variance of noise. For the τ seconds period the variance will be σ_n^2/n since there were assumed to be n frames/τ seconds. Signal to Noise Ratio (SNR) = $\frac{\text{variance of signal}}{\text{variance of noise}}$

as $\frac{\sigma_n^2}{n}$ decreases, SNR increases

and $\frac{\sigma_n^2}{n}$ decreases as n increases

Thus, for a higher n (i.e. for a higher frame rate) there will be a higher SNR

* For a stationary picture.

so that there appears to be less noise in the picture. In other words, at slow frame rates, the perceptual system "forgets" the information between frames. At high frame rates the signal is constant but the noise changes from frame to frame. When successive frames are averaged (integrated) the signal will appear to show more distinctively.

CONCLUSIONS

A first conclusion is that trained human operators could (much to their own disbelief) perform fairly complicated remote manipulation tasks with a coarse, intermittent digitized picture requiring as little as 50,000 bits/second.

Further, for a picture at 128x128 resolution, 28 f/s frame rate and 4 bits of grayscale, each of these three parameters could be decreased considerably individually, without preventing the operator from accomplishing the task.

In the range of operation (up to R = 128x128, F = 28 f/x, G = 4 bits) frame rate and grayscale could be degraded by greater factors than resolution before making task accomplishment impossible.

For the given tasks and manipulator, "threshold points" existed for all three parameters:

For FRAME RATE: 3 f/s when resolution = 128x128, grayscale = 4 bits

For RESOLUTION: 64x64 when frame rate = 28 f/s, grayscale = 4 bits

For GRAYSCALE: 1 bit

Any further degradation of a parameter beyond these points (while holding the other two constant) resulted in degradation in performance such that the task could not be completed.

It was also observed that lowering the sampling rate created more problems than making the display slower. There appeared to be more noise in images at low frame rates. This was believed to be because of the low pass nature of the visual nervous system.

A final general observation is that as F, R and G were varied, performance tended to correlate very well with bits per second in the picture.

REFERENCES

1. Ranadive, V., Video Resolution, Frame Rate and Grayscale Tradeoffs under Limited Bandwidth for Undersea Teleoperation, MIT SM Thesis, Sept. 1979.

2. Cole, R.E. and B.H. Kishimoto, Remote Operator Performance Using Bandwidth-Limited TV Displays: A Review and Proposal, Naval Ocean Systems Center Rep. NOSC-TD-379, August, 1981.

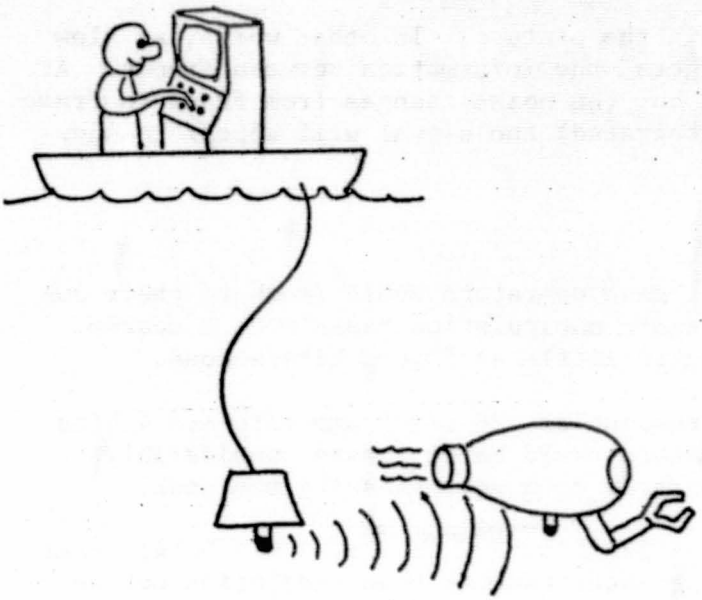


Figure 1. Undersea telemanipulator with acoustic telemetry

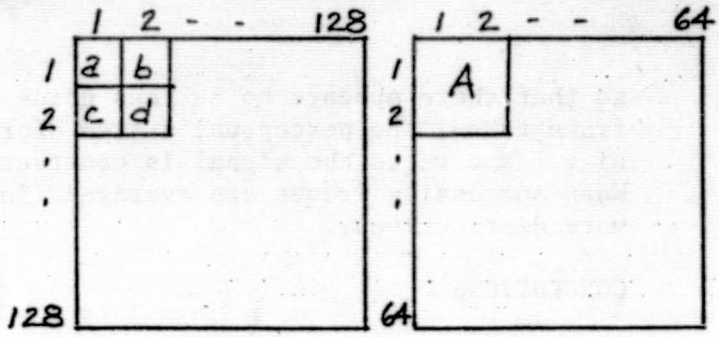


Figure 2. Resolution averaging

Figure 3. The same picture at various combinations of resolution and grayscale

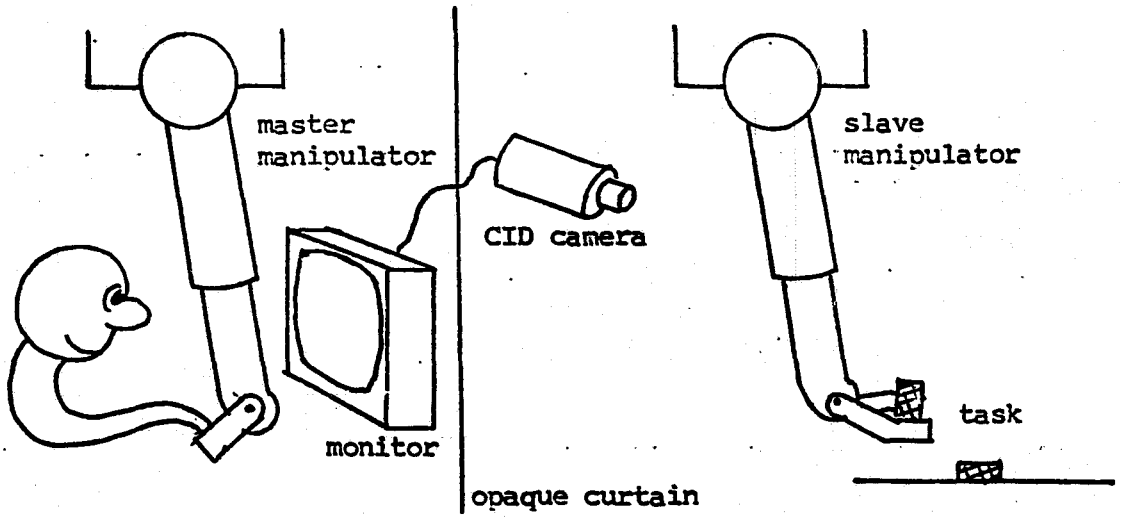


Figure 4. Experimental configuration

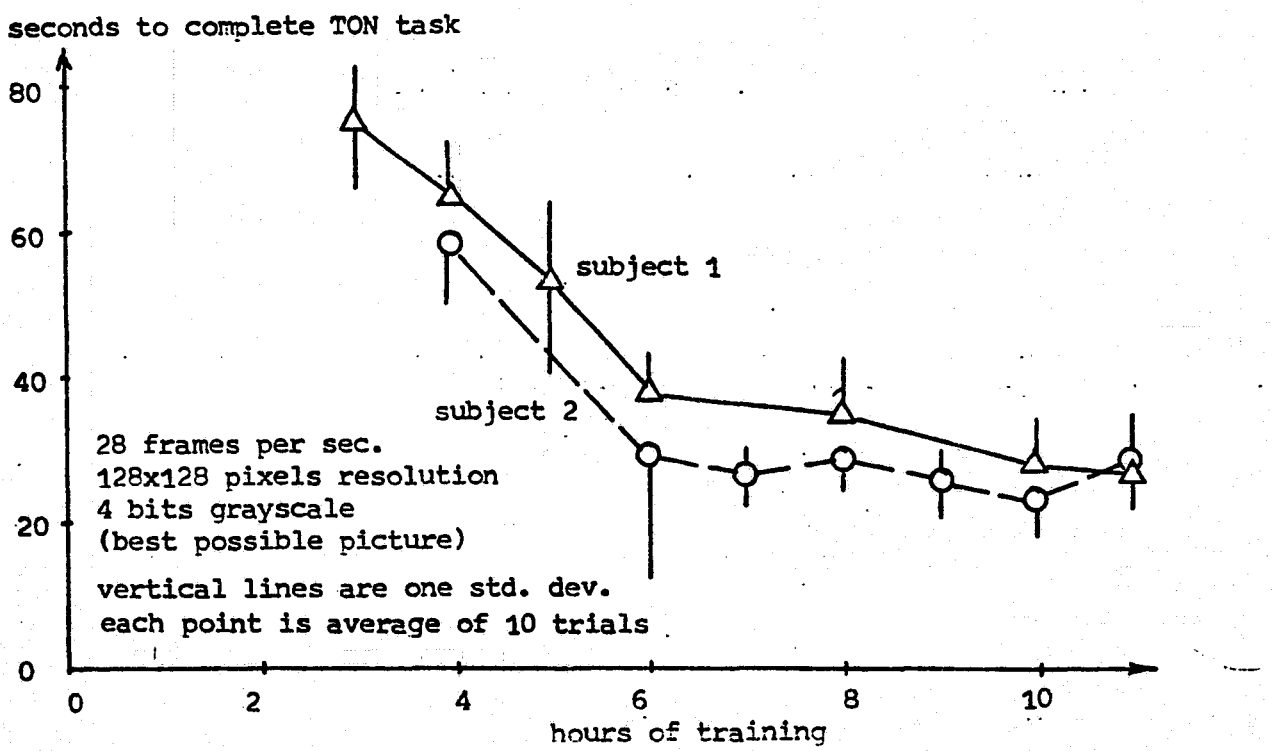


Figure 5. Learning curves for two subjects on TON task

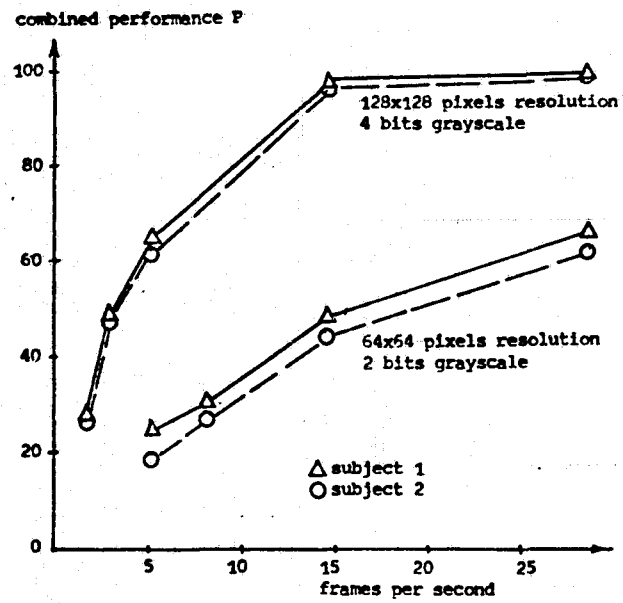


Figure 6. Effect of framerate

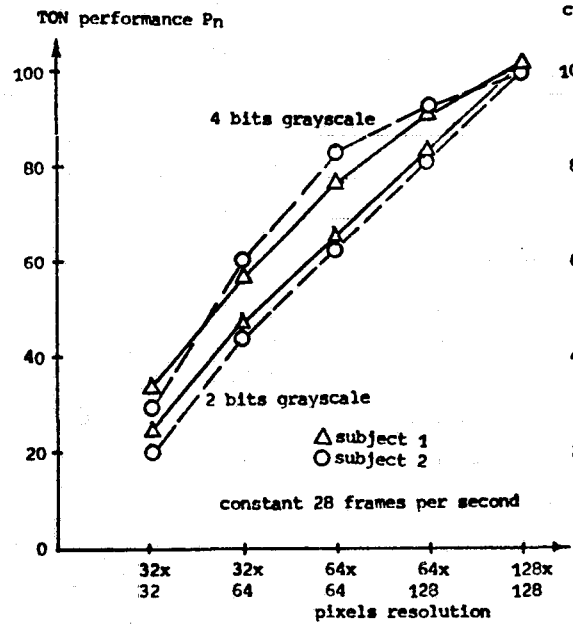


Figure 7. Effect of resolution

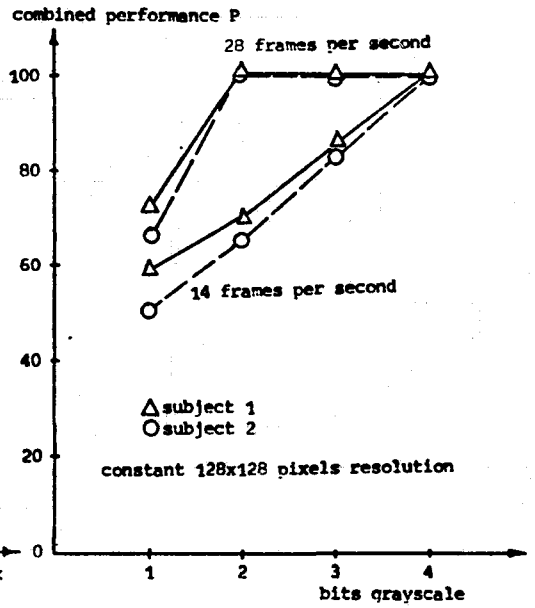


Figure 8. Effect of grayscale

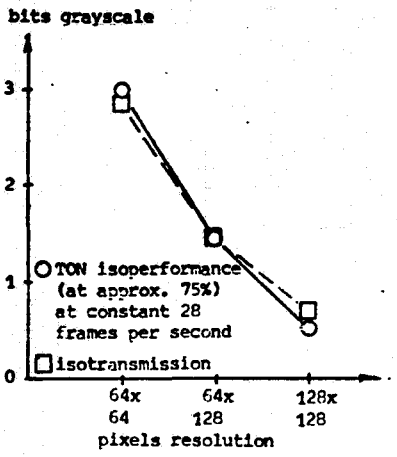


Figure 9. Isoperformance

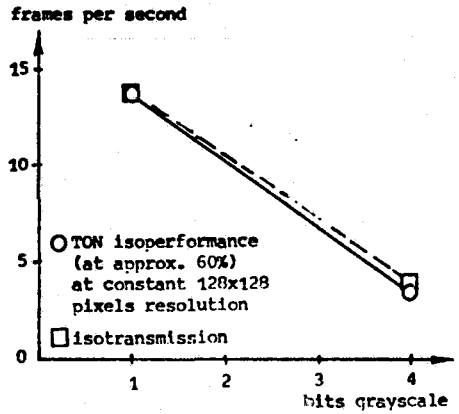


Figure 10. Isoperformance

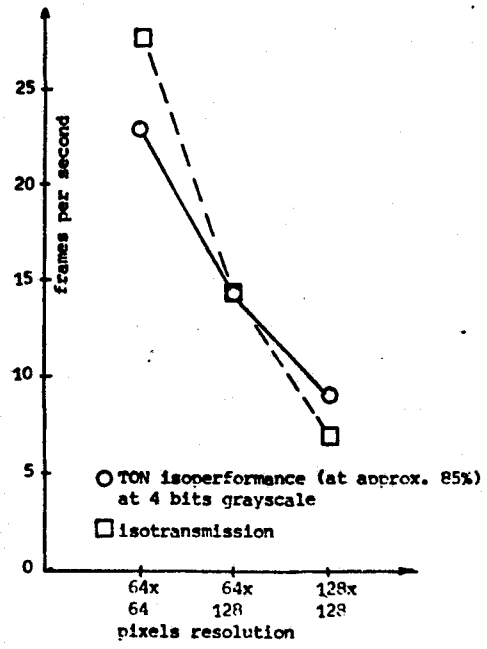


Figure 11. Isoperformance

PREDICTOR SYMBOLOGY IN COMPUTER-GENERATED PERSPECTIVE DISPLAYS

By Arthur J. Grunwald

Technion - Israel Institute of Technology *

SUMMARY

An advanced display format for the four-dimensional commercial aircraft approach-to-landing is evaluated. The desired curved and descending approach path is presented by displaying the perspective image of a tunnel. Attention is focussed on the predictor symbology, superimposed on the tunnel image. A perspective three-dimensional predictor symbol, providing future position, as well as future attitude information, is compared with a flat two-dimensional version, which only provides the future position. In addition to this, the predictor displays the actual airspeed as well as the desired airspeed, prescribed by the four-dimensional path.

Results show that the three-dimensional predictor symbol outperforms the two-dimensional predictor in following the trajectory in a moderate-to-heavy turbulent environment, which is manifested in a significantly lower roll-acitivity and a better following accuracy. Futhermore, accurate manual true airspeed control was obtained without affecting the main task performance significantly.

INTRODUCTION

Computer-generated pictorial displays facilitate the integration of control information in a format, analog to the "through-the-windshield" visual field. The tunnel display, in which the three-dimensional approach path is displayed as a winding descending "tunnel-in-the-sky" is found to be suitable in particular for following complicated curved trajectories. It is shown in a previous work [1], that pictorial displays without further augmentation, yield impaired system damping due to the lack of peripheral visual cues. It is also shown that superimposed predictor symbology furnishes the system with the necessary damping cues.

---

* This research is carried out under NASA Contract NASW-3302. Contract Technical Director: Samuel A. Morello, Flight Systems Branch, Langley Research Center, Hampton, Va.

In this paper the effectiveness of more complex predictive information is explored. A perspective aircraft symbol is shown, predicting the vehicle position as well as vehicle attitude angles, at a given time in the future. Forward velocity cues are derived from predictor-distance variations. Since, for a fixed prediction time, the predictor-distance is proportional to the vehicle velocity, the predictor symbol image will apparently shrink with increased velocity and grow with decreased velocity. The usefulness of these changes in predictor distance for controlling the true airspeed, is investigated.

DISPLAY FORMAT

Fig. 1a shows the tunnel display with perspective predictor symbol. The vehicle is to the right of the tunnel and banked to the left, as shown by the inclined horizon (a). The curved tunnel trajectory, with a square 300 x 300 ft. cross-section, is indicated by the four corner lines (b). The solid square (c) is a cross section of the tunnel a distance $D_o = T/V_o$ ahead of the vehicle, (in Fig. 1a $D_o = 900$ ft), where T is the prediction time and V_o the nominal desired true airspeed. The predictor symbol (d) is at distance D ahead where D is predicted from the actual vehicle velocity. The corresponding tunnel cross-section at D is indicated by the four bright blinking tick-marks (e). Since the solid square corresponds to the desired velocity and the tick-marks to the actual velocity, the velocity is controlled by matching the tick-marks to the square. Fig. 1b shows the display with two-dimensional predictor at a nominal distance of 2000 ft.

RESULTS

The dynamics of a control-augmented DC-8 aircraft were simulated, with either automatically or manually controlled throttle. The nominal true airspeed was 243 ft/sec. The control task was to follow the trajectory in the presence of random appearing gust disturbances. Fig. 2 shows typical results of one of four subjects. Fig. 2a shows that, both for the three-dimensional as well as for the two-dimensional predictor, the covariance of the lateral deviation increases with the nominal predictor-distance. However, the three-dimensional predictor yields a better tunnel following accuracy. Fig. 2b shows a general decrease in the covariance of the roll-rate with increased predictor distance. The advantage of the three-dimensional predictor is clearly demonstrated by the significantly lower roll-activity. In Fig. 2 the results for auto and manual throttle are compared. Manual velocity control was accomplished without affecting the following accuracy or roll-activity, significantly.

CONCLUDING REMARKS

Predicted attitude information, provided by the more complex three-dimensional predictor symbol, is successfully utilized and contributes to improved system damping. Airspeed control by using variations in the predictor distance is proved successful. The subject of further research will be the choice of predictor law and filters considering noisy on-board measured sensor data.

REFERENCES

1. Grunwald, Arthur J., Robertson, James B., and Hatfield, Jack J.: Experimental Evaluation of A Perspective Tunnel Display for Three-Dimensional Helicopter Approaches. AIAA Journal of Guidance and Control, Scheduled for Publication, Oct. 1981.

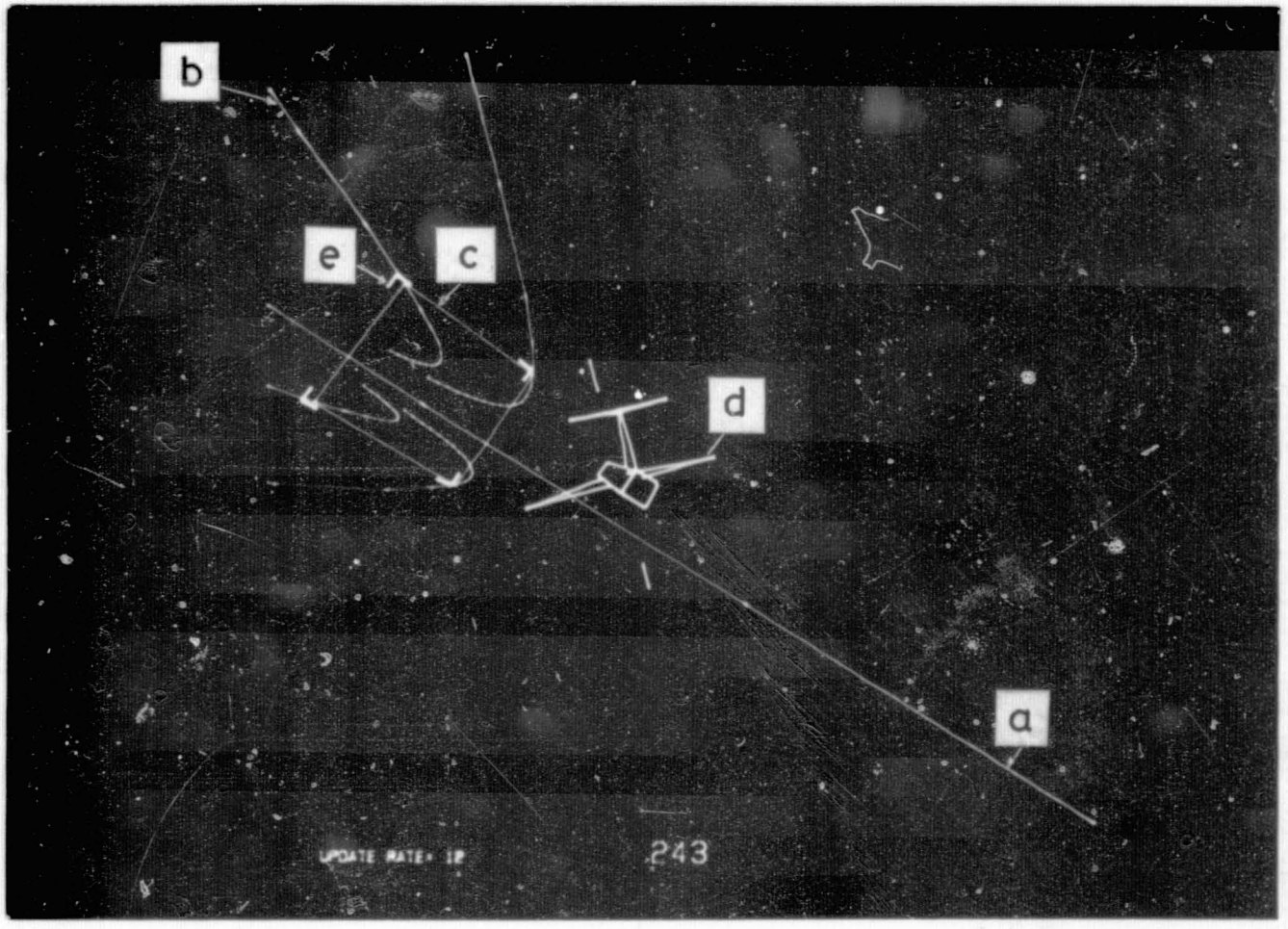


Fig. 1a: Tunnel Display with Three-Dimensional Predictor Symbol (D_o = 900 ft)

(a) Horizon Bar

(b) Corner Lines

(c) "Desired" True Airspeed Square

(d) Predictor Symbol

(e) "Actual" True Airspeed Tick-Marks.

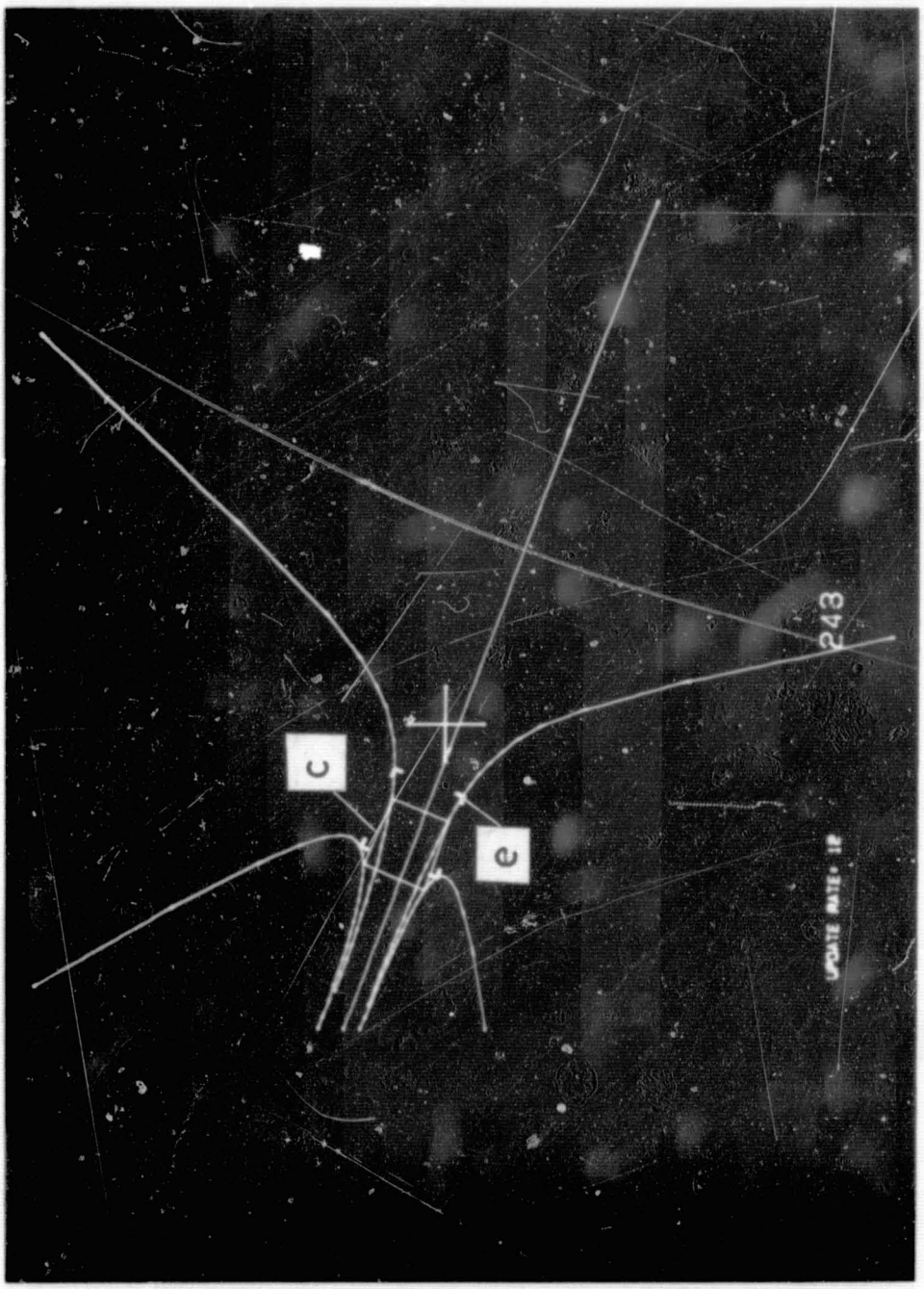


Fig. 1b: Tunnel Display with Two-Dimensional Predictor Symbol (D_o = 2000 ft).

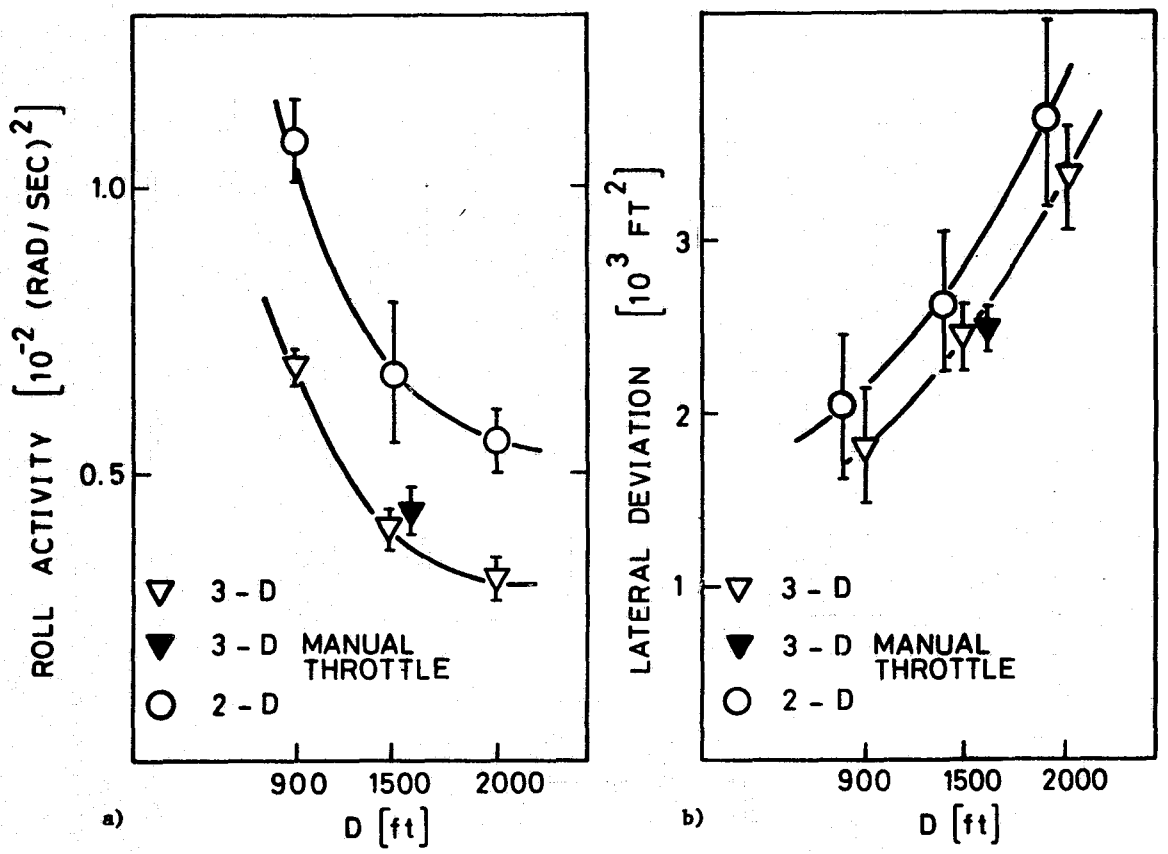


Fig. 2: Results for Tunnel Following

a) Covariance of Roll-Rate

b) Covariance of Lateral Deviation.

Average and $\pm 1\sigma$ bars of 8 runs of 128 seconds duration.

D/6

SUPERVISORY CONTROL OF REMOTE MANIPULATION: A PRELIMINARY EVALUATION

Gregory P. Starr

Department of Mechanical Engineering
The University of New Mexico
Albuquerque, New Mexico 87131 USA

ABSTRACT

Supervisory control, where control is traded between man and computer, may offer benefits in the control of a remote manipulator. A system for the study of supervisory control is described, and some preliminary results presented.

I. INTRODUCTION

Man's technological rise has been accompanied by an increasing need for him to work in hostile environments. Nuclear waste disposal sites, radioactive laboratories, the depths of the ocean, the vacuum of outer space, and underground mines are examples of such environments. Teleoperator systems project man's manipulatory capabilities into the remote environment, allowing his functional presence without his physical presence.

The need for teleoperators implies operating conditions which exclude or impair visual or other human sensory contact between the operator and the manipulator. The barrier imposed by the hazardous environment carries with it limits on both sensor and control communication.

It is possible for a man and computer to cooperate in the control of a remote manipulator, and achieve a standard of performance beyond that possible by either alone. This mode of control has been termed supervisory control [1].

The first manipulators were master-slaves used for radioactive materials handling. They provided bilateral force feedback, and the operator viewed the site directly a few feet away. In this environment the operator was capable of precise positioning and fine control of applied forces. However, as the linkage between human operator and manipulator was made electrical, and the distance increased, the impairment of sensory feedback made control more difficult.

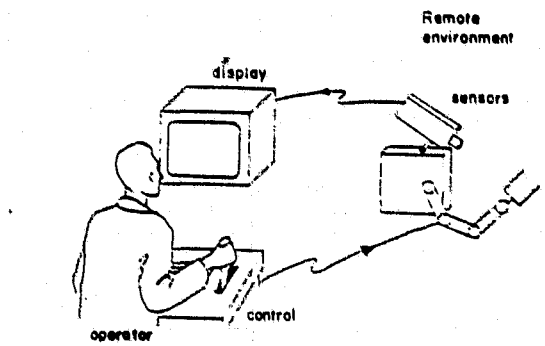
The first sensory degradation that was investigated was a transmission

time-delay between the operator's command and the manipulator's response. A time-delay can arise from extreme separation (outer space manipulation), or sensor/display limitations such as limited video frame rate.

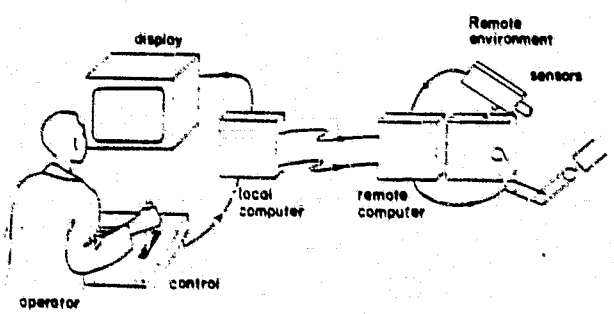
Ferrell [2], Black [3], Hill and Sword [4], and Starr [5], experimentally verified the negative effects of time-delay on manipulation. Simple tasks became frustrating and laborious, and manipulation requiring high accuracy was virtually impossible.

Supervisory control may allow manipulation to be performed effectively even when a time-delay is present. The philosophy of supervisory control is that the human operator plans strategy, monitors performance, and intervenes when necessary, while the computer accomplishes portions of the task as instructed by the human. Supervisory control allocates control responsibility between man and computer such that the inherent attributes of each are used to best advantage.

Figure 1 shows diagrams of manual control and supervisory control of a remote manipulator system. Under manual control the operator's commands are sent directly to the manipulator, and feedback from its sensors is displayed directly to him. Under supervisory control the operator's commands are transmitted to a remote computer, which then commands the manipulator. The remote computer also processes the sensory information from the manipulator and relays it to the operator. Since the remote computer is "on site" it is



MANUAL CONTROL



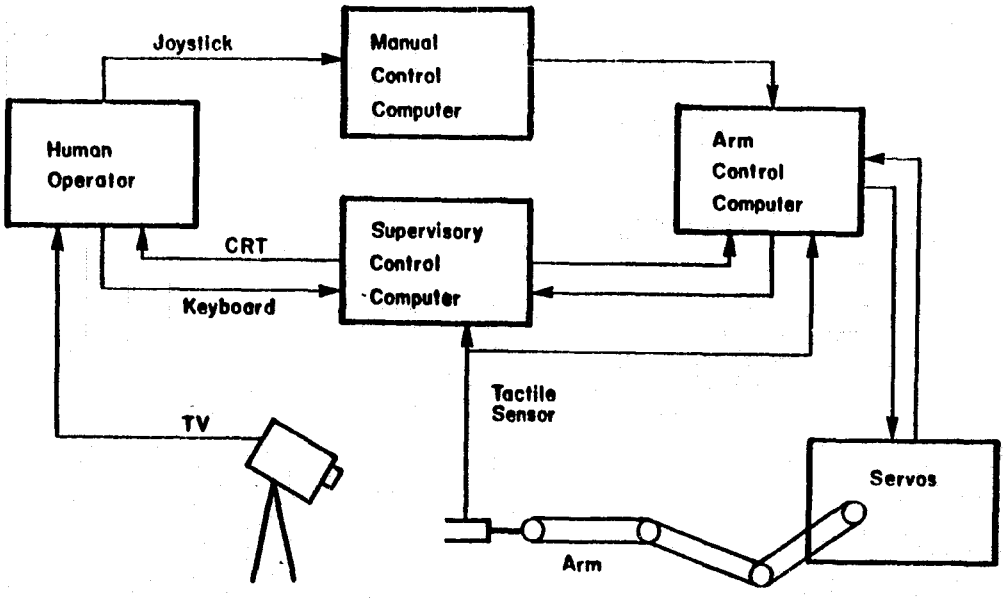
SUPERVISORY CONTROL

Figure 1

not subject to the same feedback degradation as the human operator. The portion of the system comprised by the manipulator and remote computer can accomplish portions of the task on its own. The human functions principally as coordinator and supervisor, but may still control the manipulator directly if necessary.

II. SYSTEM OVERVIEW

Figure 2 shows a block diagram of the system. There are three separate digital computers, each with its own responsibilities. The Manual Control Computer (MCC) processes manual control signals originating from the human operator. The Arm Control Computer (ACC) performs arm trajectory calculation, monitors the jaw-mounted tactile sensors, and drives the arm through



BLOCK DIAGRAM OF SUPERVISORY CONTROL SYSTEM

Figure 2

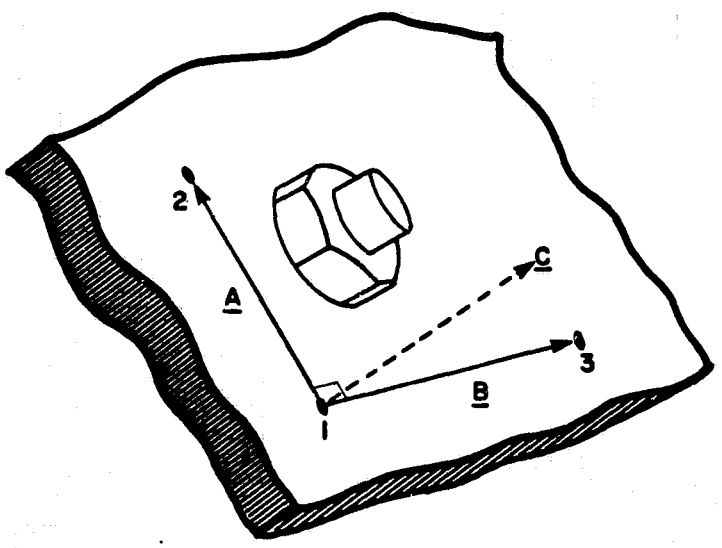
desired positions. The Supervisory Control Computer (SCC) monitors the sensors, sends commands to the ACC, and performs quantitative analysis of arm positions and sensor data.

The SCC and ACC are coupled by a 9600 band serial communications link. The ACC executes the VAL language while the SCC uses the RT-11 operating system, with programs written in FORTRAN or assembler language.

Example of System Operation. Although the individual elements in the system will be described later, the following alignment routine (used in the experiments) is an example of how they work together.

Consider the operation of removing a nut from a stud, when the plane on which the stud is fastened is arbitrarily oriented and the manipulator hand is "far away" from the nut. Before a pre-programmed nut-removal routine can be executed, the nut must be approached, and the hand must be "lined up" with the axis of the nut, i.e., normal to the plate. The sequence of operations is as follows (see Figure 3):

1. The human operator brings the hand to the vicinity of the nut, manually orients the hand to be roughly normal to the plate, then instructs the SCC to execute the "ALIGN" routine. The SCC commands the ACC to execute its stored alignment routine. The ACC advances the hand in the direction it is pointing until it touches the plate at location 1.



1. MAN MOVES ARM TO TOUCH AT "1"
2. COMPUTER MOVES ARM TO TOUCH AT "2" AND "3"
3. VECTOR $\underline{C} = \underline{A} \times \underline{B}$ COMPUTED
4. COMPUTER ORIENTS HAND ALONG $\underline{C}$
5. CONTROL RETURNED TO MAN

ALIGNING HAND WITH BOLT AXIS

Figure 3

2. The ACC detects the touch, stores the position where the touch occurred, and moves the hand to touch the plate at locations 2 and 3, storing these locations.

3. The three locations define vectors A and B, from which vector C is computed by the ACC.

4. The ACC drives the hand to point along C. The hand is now aligned with the nut axis.

5. The SCC notifies the human operator that the hand is aligned, and displays the options for continuation, which may include resumption of manual control, or invocation of another computer routine.

The remainder of the paper will be devoted to describing the elements of the system and discussing a preliminary experiment and results.

III. MANIPULATOR

The manipulator is a PUMA 600 manufactured by Unimation, Inc. The arm, shown in Figure 4, has six revolute joints, each powered by a DC servomotor. The arm has a position repeatibility of 0.1 mm (0.004 in.). It can apply a static force of 58.0 N (13.0 lb). Its length is approximately one meter.

Each arm joint is under the local control of an R6503 8-bit microprocessor, which receives commands from the Arm Control Computer (ACC), a DEC LSI-11. Each 6503 has its own PROM, RAM, and I/O logic, and servos its joint via a D/A converter and shaft encoder.

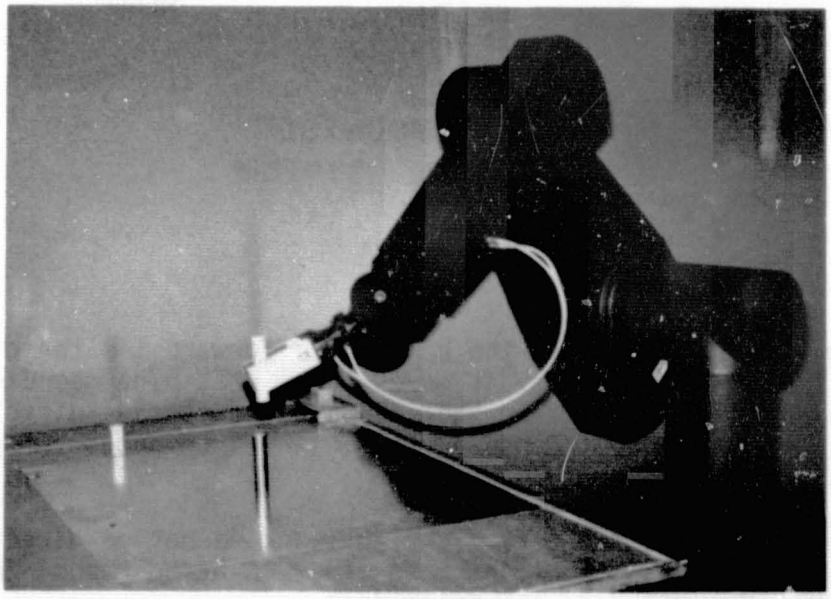
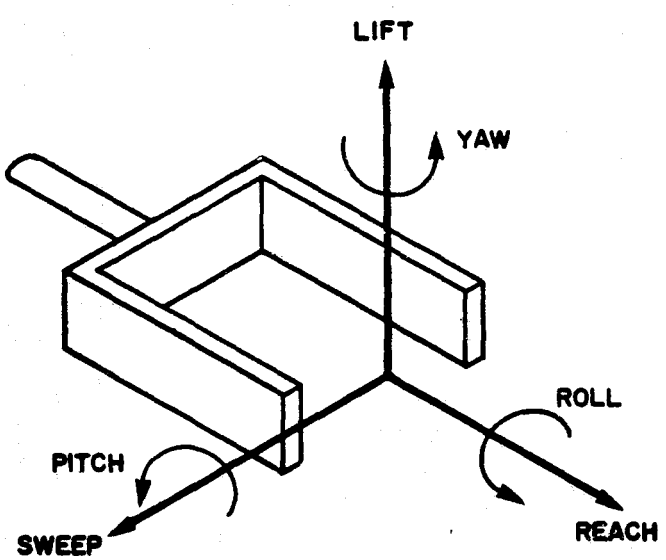


Figure 4

IV. MANUAL CONTROL SYSTEM

Resolved motion rate control (RMRC) is the manual control mode. With RMRC, the operator uses a six degree-of-freedom isometric joystick to specify velocity components of the hand along axes of a hand-mounted cartesian coordinate system. The hand-mounted coordinate system is shown in Figure 5.

The six degree-of-freedom isometric joystick, shown in Figure 6, produces six output signals proportional to the six applied forces and torques. The magnitude of a force/torque component determines the value of corresponding translational/rotational velocity component. Thus, if the operator

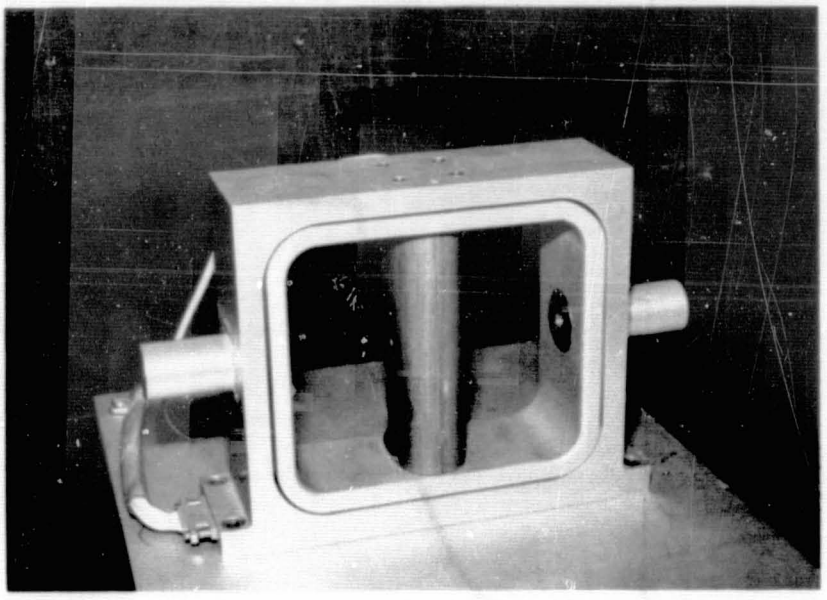


HAND – MOUNTED COORDINATE SYSTEM

Figure 5

pushes forward on the joystick, the hand moves forward (in the direction it is pointing); if he pushes down, it moves down, etc. The coordinate transformations necessary to drive the hand along the cartesian directions are performed by the Arm Control Computer (ACC), which computes arm positions each 28 msec.

The six velocity commands from the joystick do not go directly to the ACC, but are first processed by the Motorola 6800-based Manual Control Computer (MCC). The MCC does the following:



6 DOF JOYSTICK

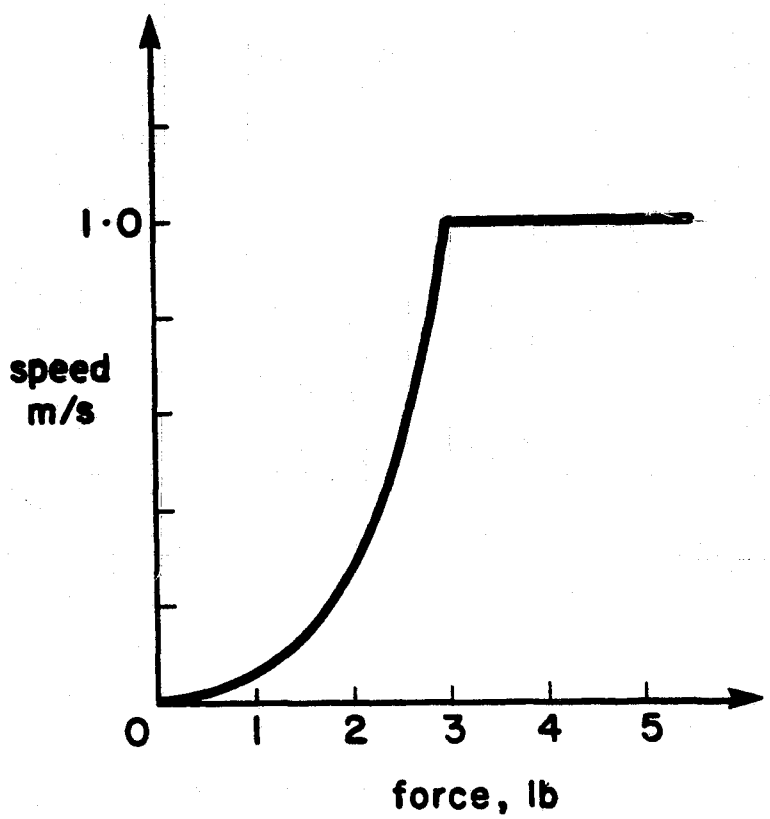
Figure 6

1. Provides nonlinear joystick force/hand velocity relationship. To get good manual control at low hand velocities, and still be able to command high speed without excessive joystick force, a cubic force/velocity relationship is implemented. Done in real time with an AMD 9511 floating point processor at a sampling rate of 50 per second, the cubic relation is shown in Figure 7. We obtain low sensitivity for low joystick force, and high sensitivity for high force. Thus, the arm can be easily driven at its minimum speed, and may still be driven at maximum speed without requiring excessively large or small joystick force.

2. Implement a time-delay. To realistically evaluate the supervisory control manipulation system, a communication barrier must be imposed between human operator and arm. The most studied, and easiest to systematically vary, is a pure time-delay. A time-delay between joystick and arm is produced by the MCC, and may be varied from zero to 10 sec, with 0.02 sec resolution.

3. Joystick deadbands. Since velocity control is an integrating process, any hysteretic joystick signal after the operator releases the joystick will cause undesired arm motion. Software deadbands in the MCC prevent this.

4. Computer control takeover. When the operator wants to switch to computer control mode, he depresses a button. On receipt of this signal, the MCC immediately switches the ACC to computer mode (from manual mode),



HAND VELOCITY VERSUS JOYSTICK FORCE

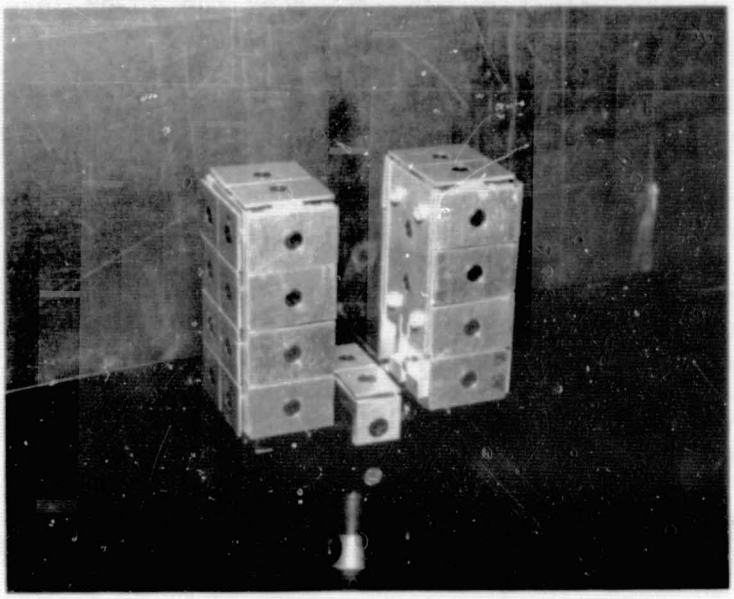
Figure 7

then after 0.05 sec the MCC issues a command to the SCC. This causes the SCC to either prompt the operator for a command, or begin a preprogrammed activity in conjunction with the ACC. The 0.05 sec wait is necessary to allow for mode-switching time in the ACC.

Note that even though this control mode is denoted manual control, it is actually a form of computer assistance where the MCC and ACC are in series with the human operator.

V. SENSOR

A jaw-mounted tactile sensor is used as the arm-based environmental sensing element. The sensor consists of a number of arrays of rectangular plates, each of which is sensitive to force. The sensor/jaw appears in Figure 8. The sensor arrays are connected to the SCC and the ACC via a multiplexer, so that when a contact is sensed, both the face of the jaw and the particular plate responsible for the contact are known.



TACTILE SENSOR JAWS

Figure 8

VI. COMPUTER CONTROL SYSTEM

The computer control system consists of two separate computers, the ACC (Arm Control Computer) and the SCC (Supervisory Control Computer). Both are DEC LSI-11 microcomputers. The ACC and the SCC work together in controlling the manipulator. Some functions performed by each machine are:

Supervisory Control Computer (SCC):

1. Can cause execution of a given ACC arm control program.
2. Read and store arm positions, numerically analyze them, take action.
3. Scan sensor, take appropriate action on contact.
4. Display prompting or status information to human operator.

Arm Control Computer (ACC):

1. Calculate trajectories, drive arm at desired speed.
2. Modify stored arm positions to be relative to new reference.
3. Read sensor, use data to govern program branching, cause jump to subroutine.
4. Execute stored routine upon request from SCC.

While some functions are common to both, the SCC controls the arm at a higher level, selecting ACC programs to executed, analyzing positions,

communicating with the operator. The ACC does the lower-level arm control, calculating trajectories and issuing joint commands.

One extremely useful capability of the ACC is that of relative positions. Preprogrammed routines can be "taught" in the laboratory, with all positions being defined relative to some reference position ("position" implies here a six-dimensional quantity including position and orientation), which is also defined at this time. When the preprogrammed routine is invoked in the field, the reference location will in general be different than that in the laboratory. The first instruction in the ACC routine can be a redefinition of the reference position to that which exists at that time. Upon execution of the routine, the ACC transforms all relative positions to be relative to the newly defined reference position. This yield the same relative hand motion regardless of the hand orientation at the time the preprogrammed routine is executed.

The ACC executes the VAL language, an interpretive language which resides in 16K EPROM. At the user level, VAL is based on common English language words, such as "move," "draw," etc., simplifying program development.

The SCC uses the RT-11 operating system, and executes programs written in FORTRAN or assembler language. Communications between the ACC and SCC, scanning of the tactile sensor, and other utility functions are done by assembler language subprograms. Quantitative processing of arm position and sensor information can be done in FORTRAN. It should be noted that the SCC and ACC work asynchronously, and while the SCC is processing data the ACC is still controlling the arm.

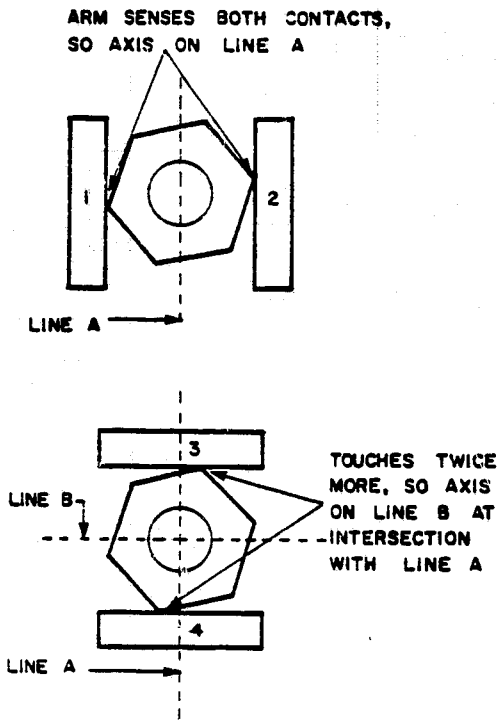
To conclude this section, the nut removal task discussed in Section II will be pursued further. In Section II, the hand was aligned with the bolt axis (Figure 3). The second phase of the task is to actually determine the location of the axis. This may be done using the tactile sensor as follows (see Figure 9):

1. After alignment, the human can instruct the SCC to execute the "LOCATE" routine. The SCC then prompts the operator to switch to computer mode.

2. The ACC moves the hand until one face of the jaw contacts the nut, stores this position, withdraws the jaw, then makes a contact on the opposite side. This defines line A.

3. The hand rotates 90 degrees, and two more contacts are made, thus defining line B. The SCC determines the intersection of A and B, thereby finding the axis.

If a time-delay is present, this task will likely be much more quickly done using supervisory control than manual control. In addition, a lower workload will be imposed on the human operator. The preliminary experiment confirmed this.



DETERMINING LOCATION OF BOLT AXIS

Figure 9

VII. PRELIMINARY EXPERIMENT

A preliminary manipulation experiment was performed, primarily to test the working of all of the system components. The experiment consisted of the removal of a nut from a bolt, and used both the "ALIGN" and "LOCATE" routines described earlier. In addition, a nut removal routine was developed to unscrew the nut from the bolt.

The operator viewed the task using a fixed, black-and-white TV camera and monitor, and was physically isolated from the manipulator by a partition. The task began with the manipulator hand positioned just at the edge of the camera field of vision, and randomly misaligned (pitch of plus/minus 20 degrees, yaw of plus/minus 20 degrees) relative to the table on which the nut was tightened. The task ended when the operator had successfully unscrewed the nut from the bolt and raised it away from the table surface a small amount.

The task was performed both with and without computer assistance, and at two time-delays, 0.0 and 1.0 seconds. Two replications at each of the four conditions yielded a total of eight trials. Only one subject participated in the experiment.

Under manual control, the subject used the six-axis joystick to control the manipulator throughout the experiment. Using computer assistance, the subject brought the hand near the table, then typed "ALIGN" to the SCC. After completion of the alighment routine, the subject brought the hand roughly alongside the nut and typed "LOCATE" to the SCC. After completion of this routine the hand was centered over the nut and ready for grasping. The subject then typed "NUT" and the manipulator unscrewed the nut under computer control. The LOCATE and NUT routines could be combined, but we preferred to leave LOCATE more general for future uses other than nut removal.

The completion times for this task are shown in Table 1. Although the sample sizes are much too small for statistical comparison, it is evident that supervisory control yielded lower completion times, especially at the 1.0 second time-delay. In addition, the subject reported a definite preference for the computer assistance.

VIII. CONCLUSION

This paper has described a system for the study of supervisory control of remote manipulation where control is traded between man and computer. By virtue of the excellent positional accuracy of the manipulator and the effective distribted processing architecture made possible by the VAL software of the Arm Control Computer and the FORTRAN/Assembler language of the Supervisory Control Computer, this system has the potential to accomplish sophisticated tasks. The human operator becomes a scene analyzer, planner, and supervisor. Preliminary experiments indicate that supervisory control with this system yield lower task completion times and is preferred by the operator over manual control.

Table 1

Time-Delay	Manual Control	Supervisory Control
0.0	163	113
	166	106
1.0	313	140
	280	125

Completion Times,
Preliminary Experiment

REFERENCES

1. T. B. Sheridan and W. R. Ferrell, "Supervisory Control of Remote Manipulation," IEEE Spectrum, Vol. 4, No. 10, Oct. 1967.

2. W. R. Ferrell, "Remote Manipulation with a Transmission Delay," NASA Technical Note D-2665, 1965.

3. J. H. Black, Jr., "Factorial Study of Remote Manipulation with Transmission Time Delay," M.S. Thesis, MIT, 1970.

4. J. W. Hill and A. J. Sword, "Study to Design and Develop a Remote Manipulator System," SRI Quarterly Report, 1971.

5. G. P. Starr, "A Comparison of Control Modes for Time-Delayed Remote Manipulation," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-9, No. 4, April 1979.

RESEARCH ISSUES IN IMPLEMENTING REMOTE PRESENCE IN TELEOPERATOR CONTROL

By Kevin Corker, Andrew H. Mishkin, and John Lyman

Engineering Systems Department, UCLA

SUMMARY

This is a position paper introducing the concept of remote presence in telemanipulation. Remote presence or telepresence is a property of an intimate man/machine interface in which the human operator is provided with a simulated sense of physical presence at the remote task site. It is suggested as an alternative to supervisory control for optimizing performance. Evidence is cited to support the contention that enhancement of a sense of presence will improve performance. A conceptual design of a prototype teleoperator system incorporating remote presence is described. The design is presented in functional terms of sensor, display, and control subsystems. The concept of an intermediate environment, in which the human operator is made to feel present, is explicated. The intermediate environment differs from the task environment due to the quantity and type of information presented to an operator and due to scaling factors protecting the operator from the hazards of the task environment. Several research issues pertaining to the development of a telepresent manipulator are delineated. The potential benefits of remote presence systems, both for manipulation and for the study of human cognition and perception are discussed.

INTRODUCTION

Remote manipulators, or teleoperators, are devices designed to allow the performance of manipulative tasks in

environments that are either too hostile or too remote to permit the physical presence of a human being. Situations in which use of remote manipulators may be appropriate include handling of highly radioactive materials and construction or exploration in space or undersea environments.

Recently efforts have been made to make the human operator (HO) of the device remote not only from the environment of the manipulation task, but also from direct control of the actual manipulations. The HO is increasingly being placed in a supervisory or "higher level" control loop (1). The trend to define the operator's function as that of supervisory controller has been based on a combination of physical constraints and assumptions concerning human performance. The HO in direct control of a remote manipulator, equipped with industry standard visual and end point displayed force reflecting feedback and vise-like grip , does not perform remote manipulation tasks as accurately or as quickly as those tasks performed by direct hands-on manipulation (2,3). The assumption is that a semi- or completely automated sensor or algorithm driven manipulator could surpass the HO performance. The physical or psychomotor constraints requiring supervisory control are ennunciated by Ferrell and Sheridan (4):

> a) the HO is so physically removed from the device as to cause inherent time delays in the control or feedback loop due to transmission lags. Such transmission lags could also be introduced in cases where a hostile environment necessitated bandwidth limitations on feed forward or feedback transmission,
> b) the operator is overburdened with other control or decision tasks.
> c) the operator is prevented from exercising direct control, either due to environmental constraints of space limitations or as a result of physical handicap, as in the case of quadriplegia.

Under many conditions physical or psychomotor constraints are not the limiting factor. If the allocation of supervisory function is based on previous performance deficits or on HO overburdening resultant from inappropriately displayed information, we contend that the operator should be placed in direct control of the remote

manipulator. Rather than removing the operator from the control loop efforts should be made to more tightly couple the HO and the manipulator.

The performance values of the HO should be exploited. The HO is adaptive and able to respond to anomaly. The operator's neuromuscular response provides for rapid, varied and fine control of manipulation. Direct control serves to reduce computational costs both in time and range of function. Finally, tasks performed in an environment whose characteristics are unstructured or unknown (and given the state of the art of computational intelligence) require that the HO be relied on for control.

METHOD

We have noted that in present systems the HO displays poor performance when in direct control of a remote manipulator. This method of control is currently characterized by limited or inappropriately coded information to the HO and by a technical subsytem of limited dexterity. We contend that this deficit is a function of the method by which control is effected rather than an inherent limitation of the HO in the performance of the task.

Examination of the literature (5) has led us to the conclusion that tightening the loop between the operator and effector will provide adequate to superior performance in manipulation. The conclusion is supported by three lines of evidence.

1) As previously reviewed, the supervisory control paradigm is appropriate when a tight link between operator and effector cannot be maintained. Such an approach is necessitated by the fact that continuous control is not possible if the HO/effector time lag increases beyond the HO's reaction time (6).
2) Physiological research in neuromuscular control indicates that tight sensory motor integration is necessary for functional motor control (7, 8, 9, 10).
3) Inference from previous research indicates that

as the link between operator and manipulator is tightened performace improves. Master slave mechanical linkage with force reflecting feedback improved performance (11). Tight mechanical link with the effector improves performance (12).

We propose that the operator/effector link can be tightened through the development of high fidelity sensors, integrated displays, true master slave control incorporating multiarticulated end effectors, and appropriate transformation algorithms. Consideration of the conceptual design parameters and recommended development for each of these subsystems follows.

CONCEPTUAL DESIGN

Figure 1 represents a potential conceptual design for a telepresent manipulation system. The arm is designed to be a position feedback system with force proportional control. A review of the state of the art of the subsytems indicated has recently been made (5) and will not be detailed here. The subsystems will be reviewed with attention to the requirements of an effective man/machine interface and to the research issues which need to be addressed for system realization.

Sensor Subsystem.

The sensor subsytem of the proposed advanced manipulator system serves the dual function of sensing operator input and commands through the master arm and sensing the condition of the remote arm, including environmental influences on that arm.

Position sensors referencing the condition of the master arm are well within the state of the art. We have in Figure 2 indicated that a simple rotary potentiometer is sufficient to provide the necessary control information.

A more challenging aspect of the sensor subsystem are those sensors which are to provide information about the remote arm and the environment in which it is operating. Several types of sensors in several modalities are

suggested. Tactile information, including touch, slip, and pressure is required. Several technologies have developed which may provide the necessary capabilities. They have been reviewed in detail by Harmon (13). Strain gauges (14) have been considered. Pressure sensitive materials (15) have been tested. Hill and Bliss (16) have tested a polymer switch arrangement in manipulator control. Bejczy (17, 18) has developed many of the advanced tactile sensor systems which would be useful in manipulator control, including touch sensors, pressure sensors and directional slip sensors. This sensory information is to be used to drive a distributed and veridical display to the operator. It is suggested that the sensor density follow roughly the human sensitivity to these inputs, with highest density sensor distribution in the end effector and a significant reduction in density for the more proximal portions of the manipulator.

Information regarding the forces impinging on the manipulator and end effector must be transmitted to the operator for proper control. These forces must be sensed on the manipulator and localized in terms of their directon and magnitude. Judiciously placed strain guages, for example, are well suited to this task.

Sensor Research Issues

Sensors used to characterize a hostile environment must be shielded from that environment wihtout loss of sensitivity to changes in environmental conditions. Such selective ruggedization of sensors must be explored. Recent experience indicates that it is possible that as analogs of biological systems are developed, they will be as prone to failure, in the same situation, as their living counterparts.

Another issue for investigation is determination of the type of sensory information most useful to the operator. This issue is highly interactive with display capabilites. Our design suggests incorporation of as many sensory modalities as possible.

Display Subsystem

Of significantly greater difficulty than acquiring

tactile and proprioceptive information is effectively displaying it to the operator. Proprioceptive information should be displayed to the HO as a distributed function with torques felt about the appropriate axes of rotation. If the load is at the endpoint of the manipulator, as in hand grip, then loading should be displayed to the hand of the operator. If, on the other hand, the loading is distributed over the entire surface of the manipulator, as in large load lifting or encounter with distributed resistance, the loading must be displayed over the relevant surfaces of the operator's arm. The current procedure of localizing all feedback at the wrist imposes a cognitive load in force translation and a fatiguing physical load at the operator's wrist. An exoskeltal position feedback system is suggested to provide distributed proprioceptive information. The distributed surface pressure display (Figure 4) responds to surface sensor activity at the remote location. The display will provide distributed touch sensing ability to the operator. In addition the sleeve provides for temperature display via hydraulic chambers responsive to temperature sensors at the remote site.

Tactile information has classsically been displayed via mechanical stimulation through vibrotactile transducers. Vibrotactile stimulation has been induced by airjets, piezoelectric vibratory elements, electromagnetic and electromechanical stimulation (16, 19, 20). Electrocutaneous stimulation has been explored (21). Electrocutaneous stimulation can be used to provide sensations of pressure, pain and heat. Flexibility of stimulation parameters, portability, and increasing miniaturization potential recommend electrotactile information display.

Display Research Issues

Both electrotactile and vibrotactile stimualtion share certain research issues which constellate around questions of:

- How accurately does the sensation provided mimic natural stimulation?
- How do the percieved attributes of the sensation vary as a function of body site stimulated?
- What is the optimal display density and

distribution?

In addition to the more common display issues there are several psychophysical interactions which could be exploited in the effective presentation of electrotactile stimulation. Structural interactions provide stability and accuracy in the inherently noisy neural systems. For example, lateral inhibition in the retina is a process of local inhibition of receptors as a function of stimulation intensity. This inhibition provides a sharpening of the illumination difference among the receptors. The result is a perceptual demarcation of stimulus intensity change that has no physical corollary in the stimulus. The sharpening of tactile stimuli by the simulation of the biological transduction process of structural interaction has been initiated (22). Temporal interactions in tactile display also provide an ability to exploit physiological processing. Sensory saltation (23) and the "phi phenomenon" (24) both, through appropriate stimulation sequencing, provide stable tactile stimulation which is not associated with a particular stimulator site. These processes suggest methods of increasing display density without a cost in operator encumbrance.

Control Subsystem

The control system is intended , as illustrated in Figures 2 and 3a,b,c, to be a fully articulated master slave control. The master slave control concept is likely to maintain the tight link between operator and manipulator which is necessary to generate a sense of presence. The computational requirements in the master slave controller should be significantly lower than other types of controllers.

Control Research Issues

The integration of several interactive feedback and feedforward loops which is required in our design raises some issues asociated with control stability. It is imperative that the operator control link be maintained as tightly as possible so as to minimize control lags and to maximize system stability.

One factor which must be investigated in the control system is that of translation between an operators hand or arm motion and that of the manipulator. A disparity between these motions is inherent in the system design. For example, the hand controller not only translates the operator's motion commands to the manipulator, but also acts as a tactile and proprioceptive display to the operator. The physical requirements of the display limit the operator's range of motion. This limitation in hand closure, for instance, must be accounted for in the command structure to the manipulator.

Translation Subsystem.

The example cited above regarding control translation from somewhat limited operator movement to full manipulator motion is just one of a class of issues which we have designated as requiring a translation subsystem. The environment for which the manipulator is designed cannot be displayed directly to the operator. Similarly the control functions of the operator must be translated to appropriate patterns for manipulator kinematics. The translation is conceived to take place in an environment which is intermediate between the hostile or remote environment and the actual operator control area.

The translations must take into account:

- attenuating or filtering hostile environment influences,
- scaling sensory input to an appropriate range for the operator,
- making some modality transformations, e.g., if the manipulator is exposed to damaging influences, the information might be reasonably transmitted to the the operator as a moderately uncomfortable heat sensation which the operator could choose to eliminate if there were need to continue to operate in that environment.
- scaling the operator input to a range appropriate to the manipulator kinematics.
- compensating for system induced distortions in control or feedback, e.g., overcoming the display/control distortion in the hand controller, or compensating for system

inertial properties in the exoskeleton operator/interface.

The intermediate environment concept provides a mapping technique in which translation rules can be imposed in the feedforward and feedback loops of the manipulator system. The imposition of these translations creates an intermediate environment for man/machine interaction.

CONCLUSION

We propose the concept of an advanced manipulator system in which the operator is intimately linked to the system in a state of "remoted presence" as an alternative to current trends in teleoperator research. Examination of the required elements of a telepresent system has generated research issues which must be addressed before the prototype of such a manipulative device can be constructed:

- The necessary density of sensor arrays must be investigated.
- Given use in hostile environments, sheilding of sensor arrays without unduly limiting sensitivity or manipulative ability must be developed.
- The capability of currently available tactile stimulation technology to provide simulation of touch, and the necessity of doing so, must be evaluated.
- Applicability of such physiological phenomena as sensory saltation and inhibition to enhance tactile display capability must be assessed.
- The algorithmic requirements of translating control and sensor information between the task and intermediate environments must be determined.
- The functional environment, in which the operator is to be made to feel present, is referred to as the intermediate environment. The information content of that intermediate environment required for

optimal operator performance must be defined.
- The mechanical subsystem must be made capable of the fine manipulations that high resolution sensor and display subsystems will support.
- The complexity of the control system that integrates the sensor, display,control subsystems, and translation algorithm must be examined to determine overall system stability.

The development of a telepresent manipulator system will contribute to performance in remote manipulation tasks approaching that of a truly present human operator. By immersing the operator so completely in this man-machine system, a remote system may for the first time possess the adaptability of the human being in unexplored environments and unstructured situations. If properly implemented, the master-slave configuration of the system may substantially reduce the computational requirements of the control loop over those of a supervisory control system. Since a remote presence is resultant of the combination of displays of several sensory modalities (visual , auditory, tactile, and kinesthetic) and effector capability, it may provide an avenue for the study of human cognition and perception. In the most advanced state, the sensor/display systems of a telepresent manipulator could allow selective control over the stimuli presented to an operator. Directing research toward the development of such a system may result in both enhanced capability in remote manipulation tasks, and a powerful tool for the study of human psychomotor control.

REFERENCES

1. Sheridan, T. B. and G. Johnansen, Monitoring Behavior and Supervisory Control. Plenum Press, New York, 1976.

2. Corker, K., J. Lyman, and S. Sheredos, A Preliminary Evaluation Of Remote Medical Manipulators, Bulletin of Prosthetics Research, BPR 10-32, pp. 107-134, 1979.

3. Mosher, R.S., An Electrohydraulic Bilateral Servomanipulator, Proceedings of the Eigth Hot Laboratories Equipment

Conference, AECTID-799, 1960.

4. Ferrell, W.R. and T.B. Sheridan, Supervisory Control of Remote Manipulation, IEEE Spectrum, Oct., pp.81-88 1967.

5. Corker, K., A. Mishkin and J. Lyman, Achievement of a sence of operator presence in remote manipulation, Biotechnology Laboratory Report, UCLA Engineering report ENG-8071.

6. Gilford, R.N.,and J. Lyman, Tracking Performance with Advanced and Delayed Visual Display, Human Factors, vol. 9,pp.127-132, 1967.

7. Clough,J., D. Kernell, and C. Phillips, The distribution of monosynaptic excitation from the pyramidal tract and primary spindle afferents to the motoneurons of the Baboon's hand and forearm, Journal of Physiology, vol. 198 pp. 145-166, 1968.

8. Smith, J.L., E.M. Roberts, and E. Athens, Fusimotor neuron block and voluntary movement in man, Journal of Physical Medicine, vol. 51, pp. 225-238, 1972.

9. Hepp-Reymond, M. and M. Wiezendanger, Unilateral pyramidotomy in monkeys: effect on force and speed of conditional precision grip, Brain Research, vol. 36. pp. 117-131, 1972.

10. Beck, C.and W. Chambers, Speed, accuracy, and strength of forelimb movements after unilateral pyramidotomy in Rhesus Monkey, Journal of Comparative Physiology and Psychology, vol. 70, pp. 1-22, 1970.

11. Mosher, R.S. and B. Wendell, Force Reflecting Electro-Hydraulic Servomanipulator, Electro-technology, vol.66 1960.

12. McGovern, D., Task analysis scheme with implications for supervisory control of remote manipulators, IEEE Conference on Cybernetics and Society, 1977.

13.Harmon, L.D., Touch Sensing Technology, prepublished manuscript submitted to the SME Robotics Institute, May, 1980.

14.Binford, T.O., Sensor systems for Manipulation, in E. Heer (ed.) Remotely Manned Systems, California Institute of Technology, 1973.

15. Synder,D.S. and J. St. Clair, Conductive elastomers as a sensor for industrial parts handling equipment, IEEE Transactions on Instrumentation and Measurement,IM-27 pp. 94-98, 1978.

16. Hill, J.W. and J.C. Bliss, A computer assissted teleoperator control with tactile feeedback, Proceedings of the Seventh Annual Conference on Manual Control,pp. 349-361, 1971

17. Bejczy, A., Kinesthetic and Graphic Feedback for Integrated Operator Control, Sixth Annual Control Conference on Man-Machine Interfaces for Industrial Control, 1980.

18. Bejczy, A. Sensors, Controls, and Man-machine Interface for Advanced Teleoperation, Science, vol. 208, pp. 1327-1335, 1980.

19. Saunders, F.A. and C.C. Collins, Electrical Stimulation of the Sense of Touch, Journal of Biomedical Systems, vol. 2, 1971.

20. Holmund, G.W. and C.C. Collins, An electromagnetic Tactile stimulator, Journal of Biomedical Systems, vol. 1, 1970.

21. Geldard, F.A., The Human Senses, J. Wiley and Sons, New York, 1972.

22. Clot, J., P. Rabischong, E. Peruchon, and J. Falipou, Principles and applications of the artificially sensitive skin Proceeedings of the Fifth International Symposium on External Control of Human Extremeties, pp. 211-220, 1975.

23. Geldard, F.A. Sensory Saltation, J. Wiley and Sons, New York, 1975.

24. Bekesy, George von, Sensory Inhibition. Princeton Univ. Press 1967.

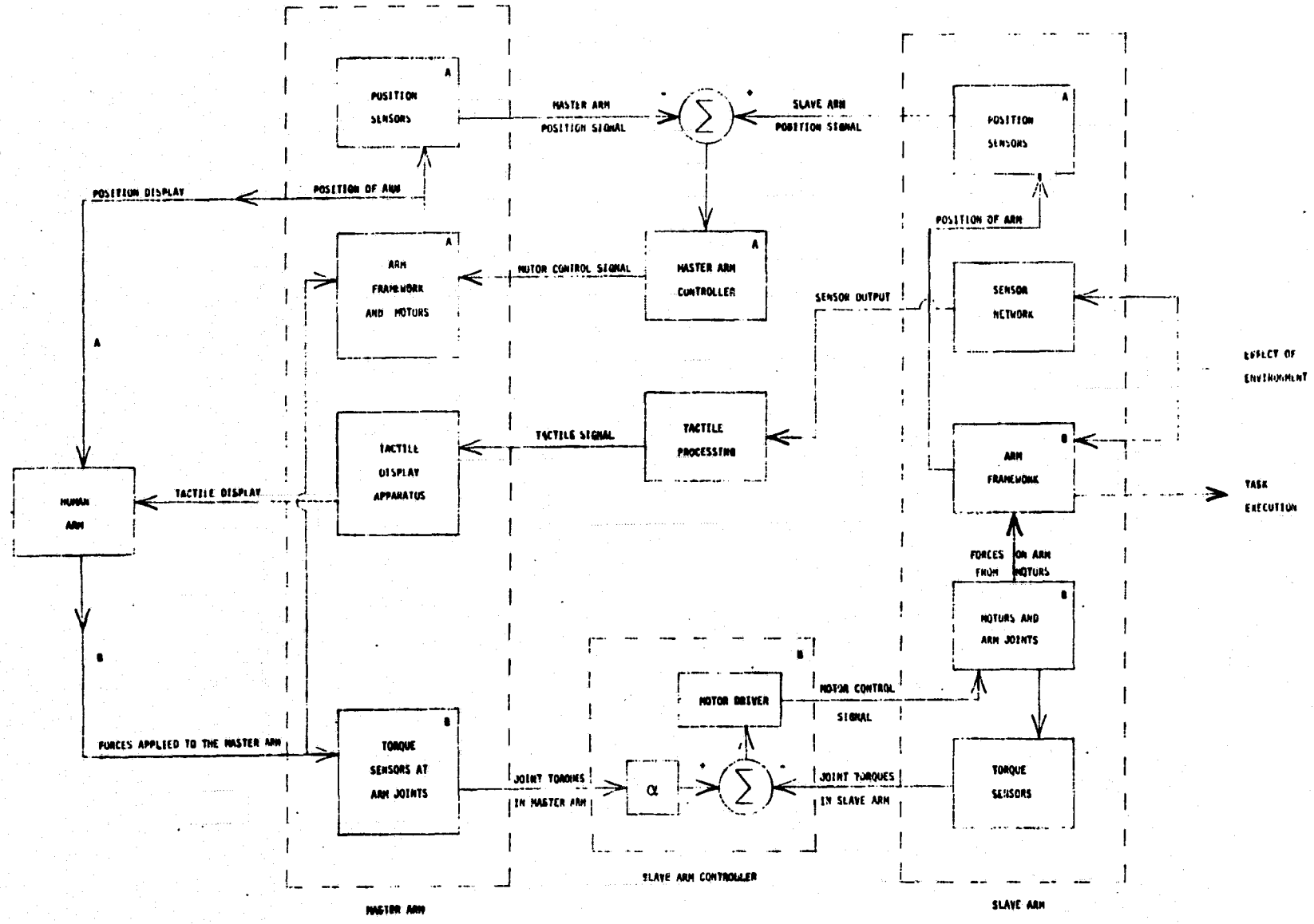


FIGURE 1 TELEPRESENT MASTER/SLAVE MANIPULATOR

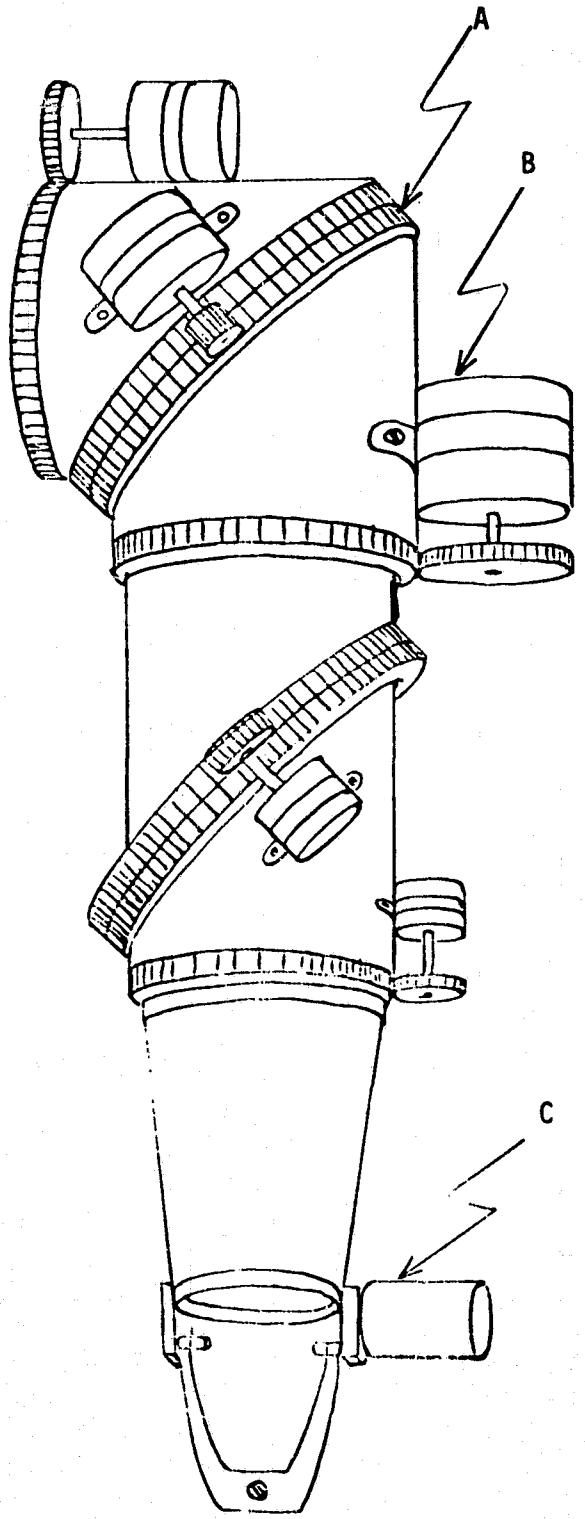


FIGURE 2 MASTER ARM

A) MASTER ARM TORSIONAL JOINT LINKAGE CONTAINING ROTARY POTENTIOMETER

B) SERVO MOTOR PROVIDING POSITION FEEDBACK

C) WRIST FLEXION-EXTENSION CONTROLLER

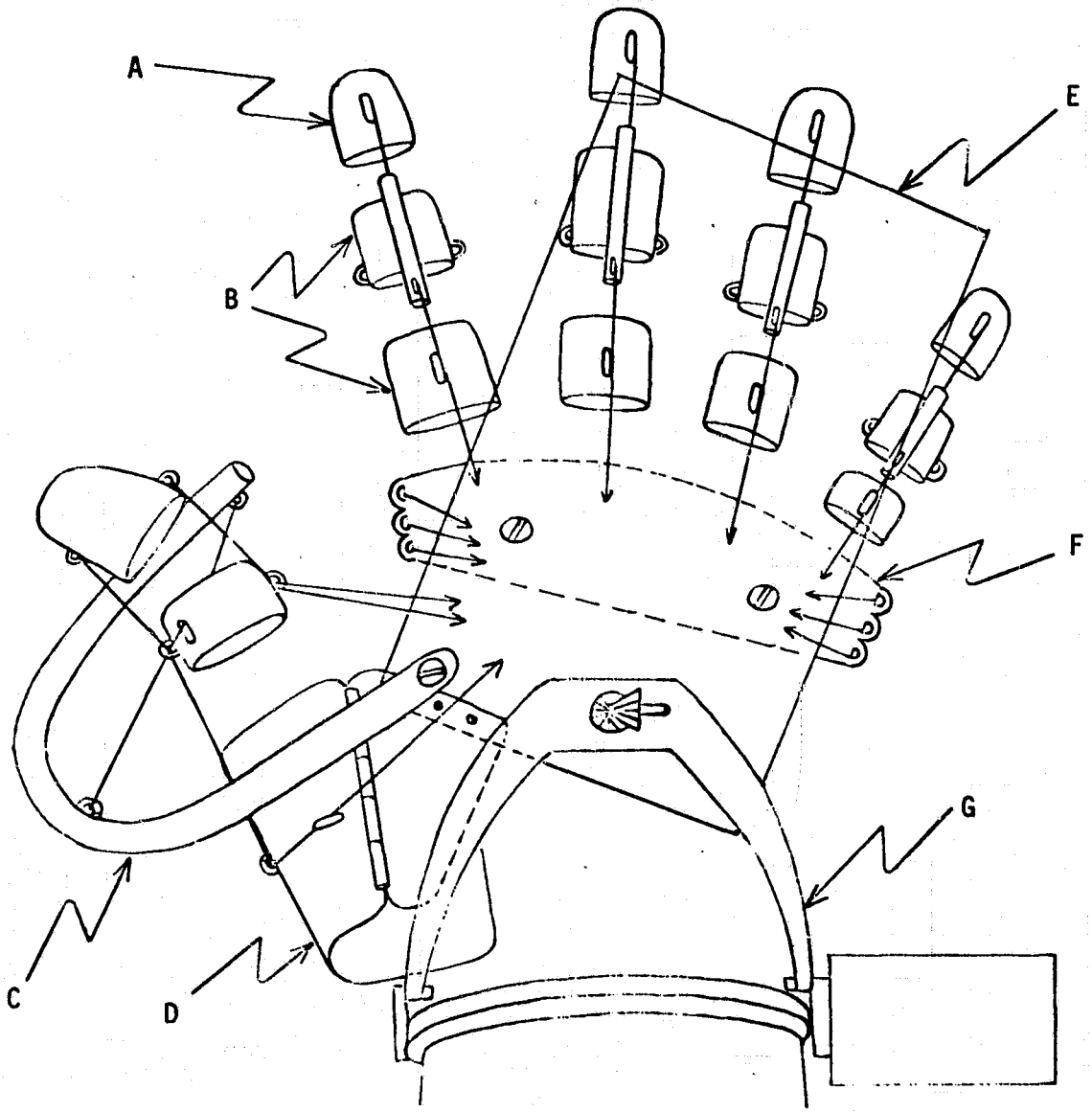


FIGURE 3a MULTIARTICULATED MASTER HAND (TOP VIEW)

A) FINGER THIMBLE AND ELECTROTACTILE DISPLAY
B) FINGER DISTROL RING
C) THUMB DISTROL OUTRIGGER TUBE
D) HINGED LOWER THUMB DISTROL
E) TORQUE MOTOR MOUNTING PLATE (MOTORS NOT SHOWN FOR CLARITY)
F) KNUCKLE BRACE
G) WRIST FLEXION-EXTENSION DISTROL

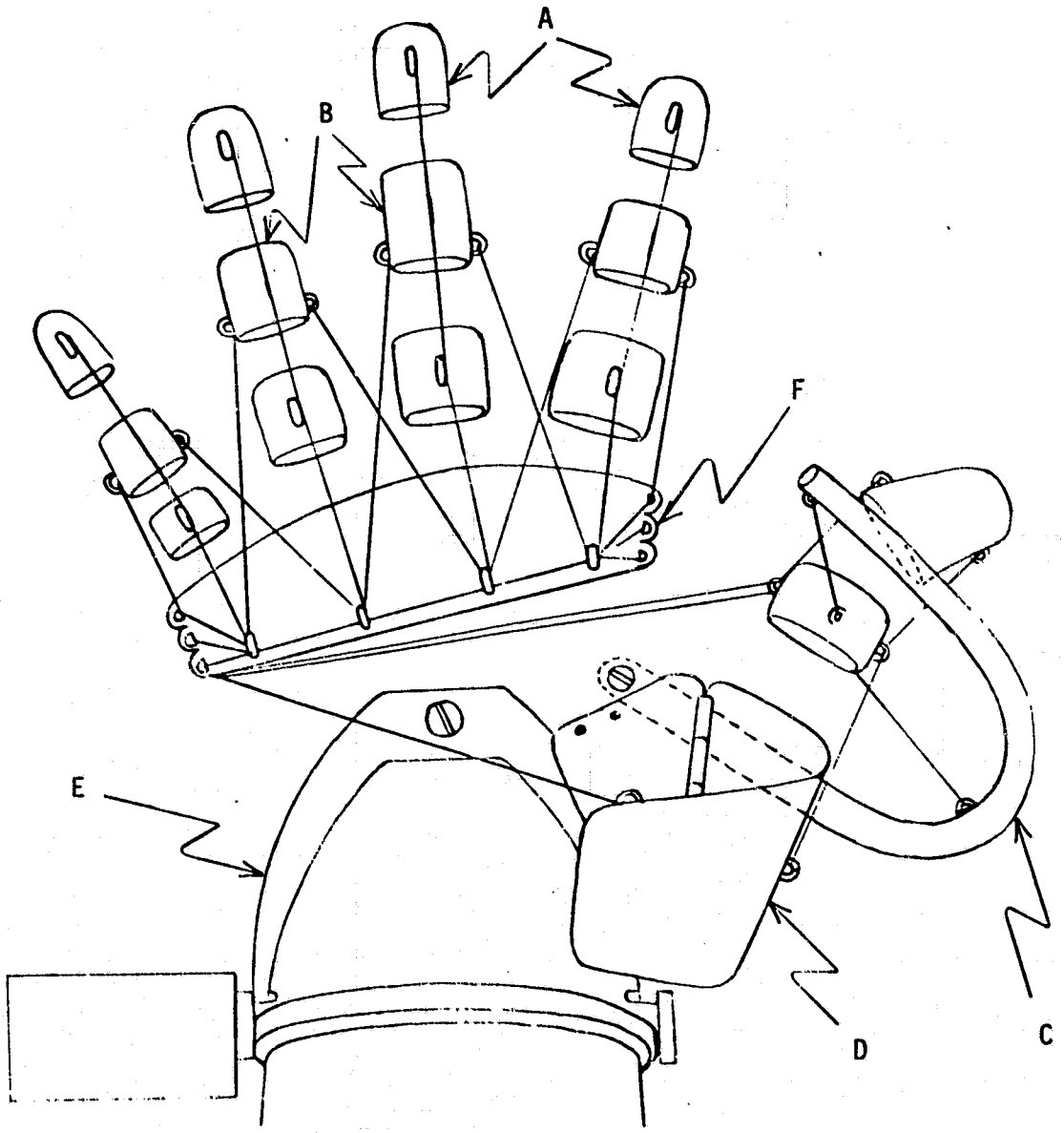


FIGURE 3b MASTER HAND CONTROLLER (BOTTOM VIEW)

A) FINGER THIMBLE DISTROL WITH ELECTROTACTILE DISPLAY
B) FINGER DISTROL RING
C) THUMB DISTROL OUTRIGGER TUBE
D) HINGED LOWER THUMB DISTROL
E) WRIST FLEXION-EXTENSION DISTROL
F) KNUCKLE BRACE

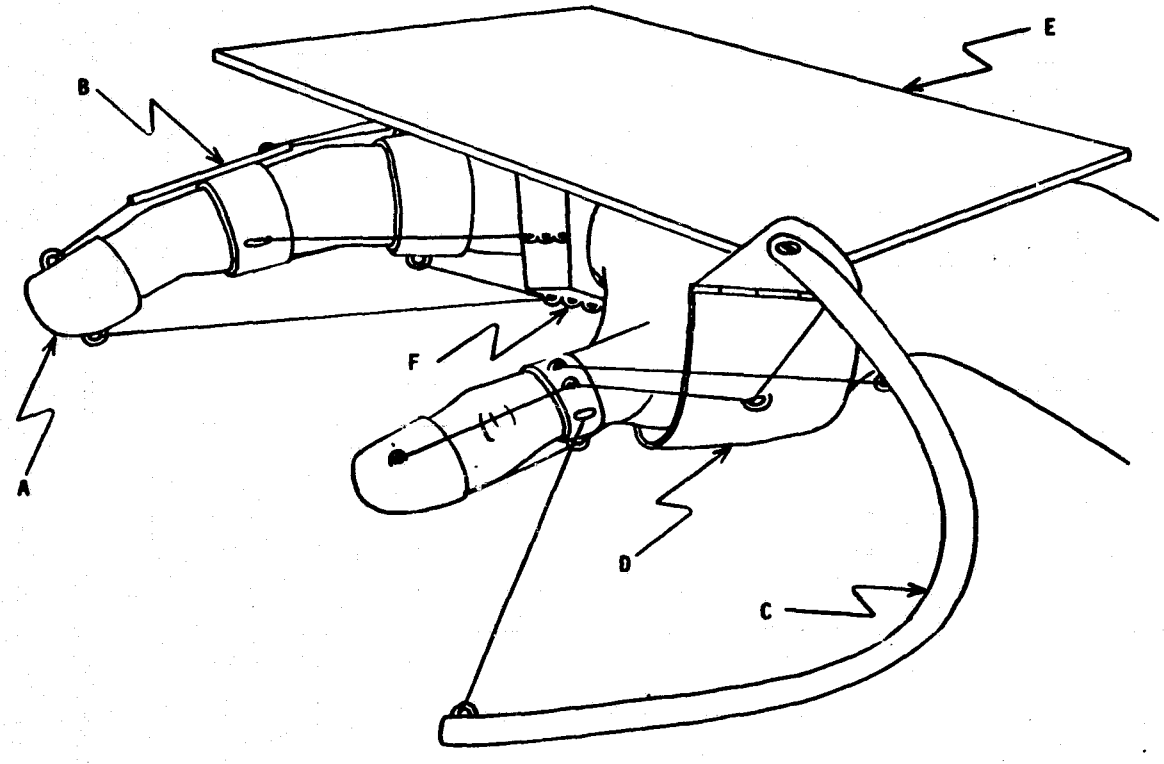


FIGURE 3c MASTER HAND CONTROLLER (SIDE VIEW)

A) FINGER THIMBLE DISTROL WITH ELECTROTACTILE DISPLAY
B) FINGER DISTROL RING WITH CABLE GUIDE
C) THUMB DISTROL OUTRIGGER TUBE
D) HINGED LOWER THUMB DISTROL
E) TORQUE MOTOR MOUNTING PLATE (MOTORS AND CABLE POSITIONS NOT SHOWN)
F) KNUCKLE BRACE WITH EYELETS AND CABLE GUIDES

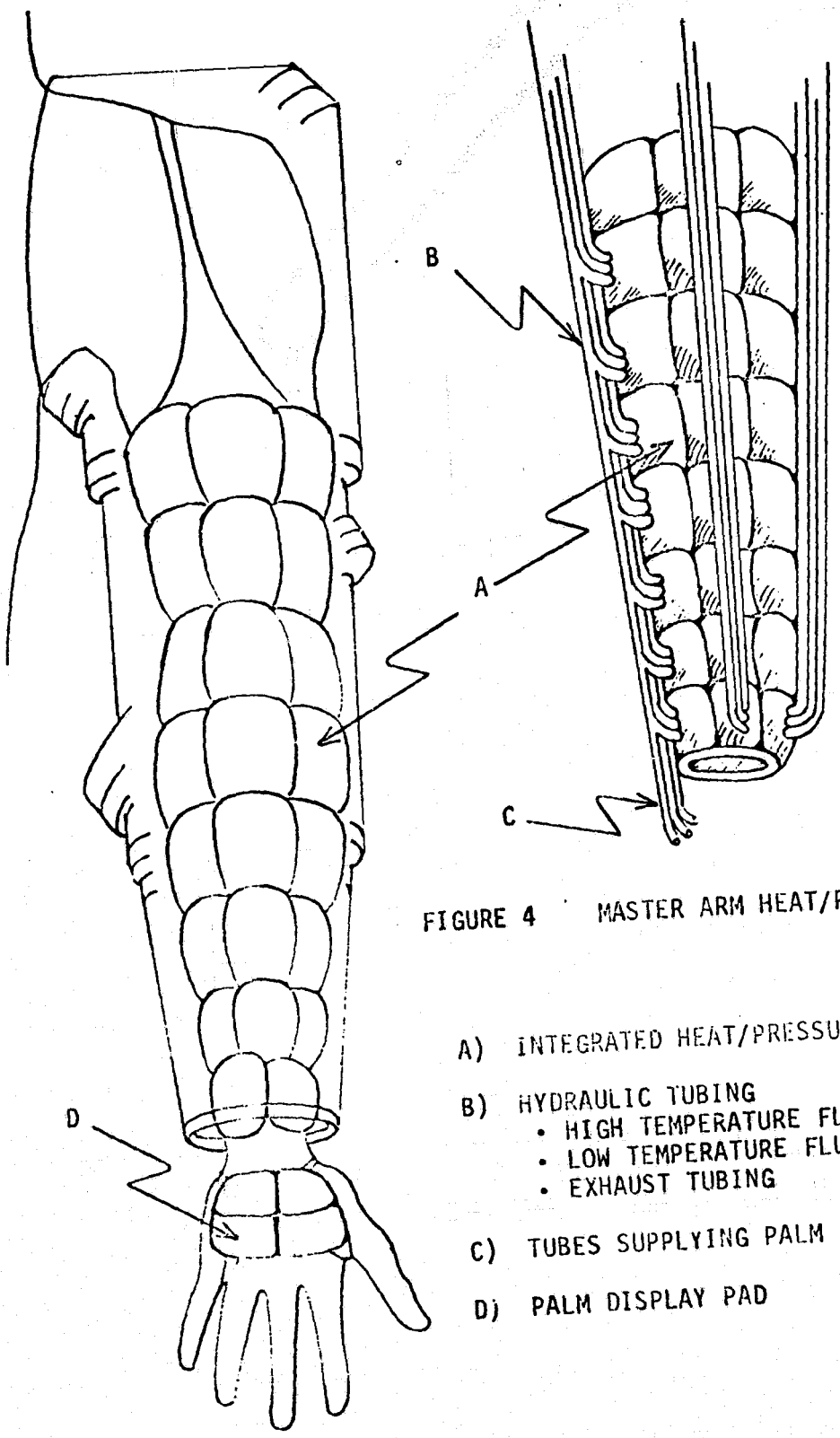


FIGURE 4 MASTER ARM HEAT/PRESSURE DISPLAY

A) INTEGRATED HEAT/PRESSURE DISPLAY POCKET

B) HYDRAULIC TUBING
- HIGH TEMPERATURE FLUID SUPPLY
- LOW TEMPERATURE FLUID SUPPLY
- EXHAUST TUBING

C) TUBES SUPPLYING PALM DISPLAY

D) PALM DISPLAY PAD

THE PERSISTENCE OF A VISUAL DOMINANCE EFFECT IN A TELEMANIPULATION TASK : A COMPARISON BETWEEN VISUAL AND ELECTROTACTILE FEEDBACK

J. P. Gaillard
Universite de Paris, XII
9100 EVRY, FRANCE

In the two experiments reported, we have examined the efficiency of an electrotactile stimulation for manipulation control. These studies were based on a psychological point of view and issued from the ergonomical purpose to feedback information to a severely handicapped operator, such as a quadriplegic, to control a telemanipulator. We have generalised the telemanipulation situation to non handicapped people operating with a hand controller and compared visual to electrotactile feedback. This sensory feedback informed operators on the sum forces exerted by the terminal device on the environment. An experimental material made of an orientable slider on a stem was developped which allowed to test the control of the three cartesian coordinates and the forces involved by the manipulator during a linear constraining movement task. Different cognitive activities were hypothesised in three situations : a first one where operators performed the task in a viewing situation, a second one where they were in a blind-folded situation but knew the spatial orientation of the stem and a third one where they had to operate blind-folded and guess the spatial orientation of the stem. Attention demand, type and quantity of information feedback were experimentaly modified so as to examine different effects such as visual dominance, alerting effect and treatment capacity. Main results showed that an electrotactile frequency modulation (1-50 Hz.) provided information which is treated as an on/off signal and is longer to access cognitive treatment and to develop operative image than visual information. Comparatively equivalent visual information is more sensible and makes decision movement faster, showing a permanent visual dominance effect even if the operator's attention is centered on the tactile modality.

In telemanipulation control evidence has been established that a force feedback system improves operator performance (Hill, 1976, Vertut and al. 1976). In the CEA - La Calhène MA 23 master-slave manipulator system the force feedback is delivred as a force reflecting the force vector on the slave handled by an operator. The Psychological interpretation for this improvement is quite simple if we assume that the force feedback presents an information increasing the operator's knowledge of the results of his control actions. Therefore movement correction and initiation can be improved by increasing such feedback to the operator. However if no force-reflecting master-slave manipulation is used, in what other may could the force information be presented ? This question is posed in particular when the master is replaced by a "syntaxer", a control transducer combining 6 degrees of freedom in a single handle.

As one possibility we studied the use of an electrotactile stimulation to return information to the operator and to test the efficiency of the electrotactile modality in teleoperator control. Several studies have been published on electrotactile stimulation, showing that a frequency modulation may be used (Shannon, 1976; Szeto, Prior and Lyman, 1977; Tachi, Tanie and Abe, 1978; Kato and al. 1970; Solomonow, Lyman and Freedy, 1977; Solomonow and Lyman, 1978). However, thus for it is not clear what cognitive processes are involved in presenting the information provided by different sensory modalities. One could argue that the electrical stimulation is transmitted to the cognitive center faster than visual or auditory stimulus. However, the attention required to code and stock information may depend on the modality through which this information accesses the cognitive treatment centers and whether parallel processing of information is possible when different modalities are stimulated simultaneously.

In Hill's studies (1974) no clear results have been obtained with a touch feedback system in telemanipulation tasks as compared to a visual feedback and he concluded that the learning effect was to strong to observe an effect on the type of display. The author recommended to alternate experimental variables in order to reduce the variance caused by learning.

Earlier, Bliss and al. demonstrated that a tactile display augmented the human response time in comparison to the one for a visual display. However, their stimulation was mechanical and for electrical stimulation the results might be different.

In manual manipulation visual and kinesthetic information is used to control movements, information is delivered bimodaly and attention is divided between two sensory modalities. Some results have shown (Treisman and Davies, 1973) that dividing attention between two different modalities is more efficient than dividing attention between signals of the same modality. We hypothesized that in teleoperator control where vision is focused on controlling the task, force feedback on a visual display presents a situation where attention is divided on the same modality (vision). In comparison, force feedback to the operator through the

tactile modality divides his attention between two different modalities and could be more efficient.

This result is explained (Richard, 1980) by assuming that the limitations of divided attention are in part caused by the conditions of the reception and storage of the sensory input.

Basically, two different models of information processing were discused. In the Broadbendt's model semantic information is treated after its selection and is limited in capacity or sequencially processed. In the Deutsch and Deutsch's model semantic analysis is processed before the selection of information which could be treated in parallel. Various intermediate models have been proposed since this two first one (See Richard, 1980 for this question). As a model we retained the hypothesis that a first kind of semantic analysis could be effectuated in a sensory memory register before transfer to the short-term-memory. Final treatment could be of a sequential nature, while surface features of the stimulus, such as physical characteristics and some semantic features, could be processed in parallel, especially in the case of perception through multiple sensory channels. If this is the case, decision and motor responses should be better in a bimodal situation than in a monomodal one.

Another argument concerns the electrical skin stimulation we used. Stevens (1938) noted that "the electrical stimulus has its principal effect directly on the sensory nerve fibers. That is, electrical stimulation bypasses the receptive structure". So if a first treatment of the information is effectuated in the receptive structure before its transfer to short-term-memory, electrical stimulation bypassing this first structure will directly access short-term-memory without a first treatment. The consequence will be a loss of information if this peripheral structure is able to initiate a first treatment of the information (see diagram 1).

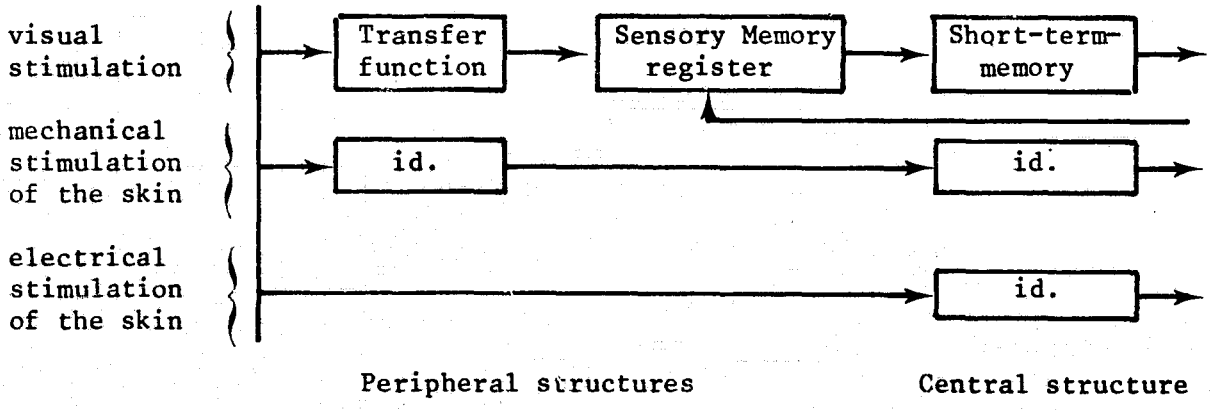


Diagram 1 : Hypothetical Peripheral Structure for visual stimulation, mechanical and electrical stimulation of the skin.

The transfer function could initiate a first treatment of the information. In the case of the visual stimulation information could be stored for a short time in a sensory memory register where information could be retrieved (Sperling, 1980).

In a previous experiment (Gaillard, 1981) testing the effect of an electrical feedback in a telemanipulation task, we did not find a significant effect on the execution time of the task. However, the number of errors with electrotactile feedback was greater than with an equivalent visual feedback. We concluded of that the gain in performance when dividing attention between two modalities could be null with electrical stimulation. Therefore, we have developped two experiments without direct vision of the task to study in more detail the role of the peripheral structure as a function of the modality of the feedback and their effects on motor control.

EXPERIMENT 1

METHOD

- Subjects : Two subjects were trained in telemanipulation control. They had been operators in previous experiments and work in the SPARTACUS Laboratory.

- Apparatus : The experimental system was composed of a manipulator type CEA - La Calhène MA 23 with six degrees of freedom plus gripper (Photo 1).

The manipulator was controlled by a "syntaxer" with six simultaneously controllable degrees of freedom. However only three of them were used for this experiment, controlling the three translations in the Cartesian coordinates. These three translations were operated with a velocity control by three micro-displacements of the handle. These micro-displacements required a certain amount of force to be exerted on the syntaxer (Photo 2).

Manipulator and syntaxer were connected through a SOLAR 16/65 computer, programmed to allow the operator to control the manipulator using a specific language linking the three degrees of freedom of the syntaxer to the three translations of the gripper in space. Initially, the system and its program were studied for a medical manipulator as a part of the SPARTACUS project as an aid to high-level quadriplegics (Guittet and al. 1979). Then the same concepts were used to control an industrial manipulator, but through another control transducer and an adapted control language.

Two types of sensory feedback were used to inform the operator about the forces exerted by the gripper on the axis. The first one was a visual display of 5 LED's with each diode successively representing a higher level of force. One diode lit meant the sum of the forces was low, but not negligible, and 5 diodes lit that the gripper was exerting a strong force. The sum of the forces was obtained from the summation of the

currents of the torque motors approximately corresponding to the three cartesian coordinates.

The second display was an electrotactile stimulation situated on the operator's arm with the following characteristics :

- rectangular pulse train
- frequency modulation 1 → 50 Hz
- pulse duration 150 µs
- pulse intensity 5 mA.

These characteristics allowed the operator to distinguish 5 levels of frequency and was sufficiently confortable. (Kato and al. 1970, Szeto, Prior and Lyman, 1977). The frequencies varied proportionaly with the sum of the forces.

- Task. The task consisted of the displacement of a cursor along an orientable axis that could be rigidly fixed with three different inclinations and twelve horizontal orientations (Photo 3). The cursor was linked to the axis by a self-aligning bearing to avoid variation of sliding friction. The task required the operator to precisely control manipulator movements in the space and to guide the cursor held by the gripper into a direction strictly parallel with the orientation of the axis.

The operator was in a "blind folded" situation. If the gripper was not guided in parallel with the axis, an increasing effort was generated and the cursor could not be conducted to the limit switch at the end stop. The sliding distance remained the same all over the experiment. The execution time was measured automatically at each trial. When the cursor began to move a microswitch was released, starting a chronometer and when the cursor had moved about 40 cms on the axis, a record microswitch was activited stopping the chronometer.

Another way to finish a trial was by the triggering of an automatic safety stop when sum of the forces exerted a certain level. Then the syntaxer was automatically uncoupled to the manipulator and the trial was considered as a failure.

- Design and Procedure. Three conditions of sensory feedback were permuted on 30 trials, 10 trials were run with visual feedback, 10 with electrotactile feedback and 10 without force feedback. The experimental design was $S_3 \times M_3 \times T_{10}$ with S for subjects, M for sensory modality and T for trials. Of the 30 trials, 21 required the control of three degrees of freedom and 9 required only two degrees of freedom. The orientation of the axis was changed for each trial and there were never two identical orientations in the 30 trials. Before each trial the experimenter fixed the axis with a new orientation and positioned the gripper on the cursor at the starting position . The operator saw the orientation of the axis and memorized it. Then, he turned his back to the axis, grasped the syntaxer, and started the trial controlling the movement only using the sensory feedback (visual or electrotactile) in 20 trials out of 30 and without

sensory feedback, only relying on his memory of orientation, in 10 trials out of 30. After each trial he was informed on his execution time of the task.

- Experimental predictions. Under electrotactile feedback condition we expected the information, bypassing the receptive structure, to be entered directly into short-term-memory (STM) where its treatment requires more time to be correctly interpreted than under the visual feedback condition where a first treatment is effectuated before its access to STM. Memory load in STM will be greater under the electrotactile condition than under the visual one and task time will be longer in the former case.

RESULTS

Task time for the 2 degrees of freedom trials were not significantly different $F(1,2) = 1,27$; $p > 10$ whereas they were different for the 3 degrees of freedom trials $F(1,6) = 8,68$; $p < .05$. No other factors, nor interaction were significantly different. Average times and number of failures were as listed in table 1.

	task time for 3 df	failure	task time for 2 df	failure
visual	13,06	2	12	0
electrotactile	17,13	0	19,93	1
no feedback	11,53	11	14,35	3

Table 1 - Task time in seconds and number of failures as a function of sensory feedback

DISCUSSION

As predicted the average time was longer with an electrotactile feedback, indicating that the task time is varying with the information treatment and memory capacities required and not with its transmission delay from peripheral to central structures. The number of failures as a function of sensory feedback showed that force feedback was of a great use to complete the task and even made it possible to execute it without feedback, only using the memorised spatial orientation of the axis. The information presented by electrotactile feedback was effectively treated and interpreted as indicated by a low number of failures and a longer task time. It suggested that electrotactile information was treated in STM and that the limit of the capacity of treatment and retention was not reached. The results also suggested that visual information could receive a first treatment in the peripheral structure before its access to STM, thereby facilitating this subsequent treatment. In this experiment feedback information had to be interpreted and compared with

the information of the memorised orientation of the axis and physical analysis of the stimulus could be sufficient to perform this comparison in memory. Now a question was raised about a semantic interpretation of electrotactile information.

EXPERIMENT 2

METHOD

- Subjects, apparatus and task remained the same as in the first experiment except for the blindfold task that had to be completed without a prior knowledge of the orientation of the axis. The subjects had to estimate the orientation of the axis during manipulation. In one half of the trials an additional task had to be completed simultaneously during the manipulation task. In this additional task, or semantic task, the operator had to listen to a text and had to press a button when he identified a proper name in the text. Subjects, therefore, were in a condition of divided attention.

- Procedure. There were four different conditions, two with attention focused on motor control with visual or electrotactile force feedback, and two with attention divided between the semantic task and the motor control with visual or electrotactile feedback. There were 40 trials for each subject, 10 for each condition. Task time, number of failures and responses in the semantic task were recorded. In addition, we recorded in X, Y recorder the sum of the forces exerted by the gripper during each trial and we asked the subject his estimation of the orientation of the axis at the end of the trial.

- Experimental predictions. To complete the telemanipulation task the subject interpreted the feedback as a confirmation or rejection of an estimated orientation of the axis, and could build a cognitive representation of the axis in space. Assuming that the electrotactile stimulation bypassed the receptive structures, we predicted that in the focused attention condition task time and failure rate would be more important with an electrotactile feedback than with a visual display as in experiment 1. If we hypothesize that a first treatment of the semantic information is made before its input in STM , then the performance will not be dramatically affected in divided attention. However, if a first treatment of semantic information is only made in STM, performance should be affected by this additional task. In this case it means that semantic information bypasses the sensory register. Indeed, we hypothesized a conscious process to interprete the sensory feedback to build an operative image of the axis in space and that process should at least involve STM storage. In other words, if semantic information is selected before its access in STM, in a preconscious activity, interpretation of the stimulus feedback should be made. If semantic information is selected in a conscious working memory, interpretation and operative image formation should be difficult. In the latter case, task time and failure rate should increase and the number of motor response decrease.

RESULTS

Mean task time and number of failures were calculated for each sensory conditions and were the following : Table 2.

Sensory conditions		Visual		Electrotactile		Visual divided		Electrotactile divided	
Dependent variables		Task time	Failures	Task time	Failures	Task time	Failures	Task time	Failures
Operators	S_1	54,2	1	86,2	22	98,1	9	115,7	16
	S_2	70,1	2	93,5	4	87,5	4	93,5	8
	m	62,1	1,5	89,8	13	92,8	6,5	104,6	12

Table 2 - Mean task time in seconds and number of failures

For the additional semantic task the percentage of alarm was 95% with visual feedback in the motor task and 96 % with electrotactile feedback. The numbers of false alarms were, respectively, 19 or 20. Guessing of the orientation of the axis was calculated from the number of degrees of freedom correctly estimated by the operator after each trial. No significant differences were found between visual and electrotactile feedback (for the two operators χ^2 were, respectively, 0.75 and 0.79 $p > .10$). The proportions of degrees of freedom correctly estimated were 72 % for the first operator and 54 % for the second one. Ranking the execution times in the condition of focused attention indicates a significant difference between visual and electrotactile (Sign test $N = 20$, $x = 1$; $p < .01$). The difference was not significant in divided attention (sign test $N = 20$, $x = 9$; $p > .10$). The only factor approaching the level of .10 in the variance analyse was the sensory condition $F(3,27) = 2,26$. In return we found a large difference between the two sensory feedbacks for the number of motor responses $t(38) = 7,82$; $p < .001$. The rate of motor response was calculated by dividing the number of responses by the task time for each trial and is indicated in table 2.

Visual	Electrotactile	Visual divided	Electrotactile divided
0,68	0,38	0,66	0,31

Table 3 - Rate of motor response

DISCUSSION

As expected results corresponded with those obtained in the first experiment. With electrotactile feedback both task execution time and number of failures increased. The score obtained in the semantic task indicated a good semantic detection when attention was divided on two different sensory channels. But the effects on task time and number of failures correspond with a decrease of performance in the telemanipulation task. Guessing the spatial orientation of the axis was more difficult with an electrotactile feedback and with attention divided, but the formation of a conscious operative image of the axis in space was necessary to complete the task and this mental image obtained with both sensory feedbacks under all conditions, as suggested by the number of degrees of freedom guessed after each trial completed.

More interesting is the number of motor responses obtained under respectively, visual and electrotactile feedback. Dividing attention had no effect on either the number of motor responses or on motor decision. Task time was increased with electrotactile because the operator could not make motor decisions during part of the time. This result suggests that the interpretation of the electrotactile feedback is more difficult and not as sensitive as of an equivalent visual feedback, and supports the general assumption that an electrotactile stimulation bypasses the peripheral structure. If so, physical features of the stimulus are not treated in this first stage and have to be treated in a working memory where the stimulus is roughly interpreted. Results also indicated that the performance in motor control and formation of a compatible operative image the axis with a visual feedback is more affected by the additional semantic task. It could be hypothesized that the semantic information is processed in a working memory after a preliminary treatment of the physical features of the stimulus in a sensory memory register. This suggests that the semantic treatment is not effectuated in parallel, but sequentially and that it is limited in capacity. With the electrotactile feedback, where no motor decision could occur for relative long periods of time, semantic information can be processed and does not interfere with the motor control.

CONCLUSION

The two experiments reported here were not designed to generalize the results to the population. Generalizations have been made for the repetition of the tasks and do only suggest direction for further research in the field of general psychology. The aim of this research was to investigate the possibility to use an electrotactile stimulation in teleoperation and to observe the interpretation of such information as a feedback to the operators. It leads us to formulate cognitive hypothesis and to propose the following assumptions :

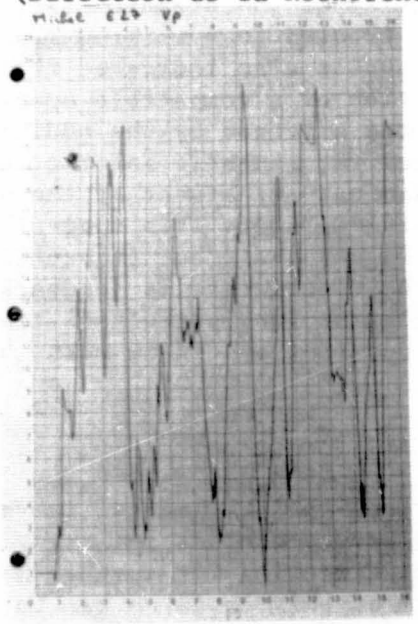
1. A visual feedback is more informative than an electrotactile one, suggested by the general visual dominance effect.

2. A complex electrotactile feedback :
   - slows down both the motor decision and motor response processes.
   - is processed as an all or nothing signal.
   - bypasses the receptive structure and accesses directly in a working memory where information is to be sequentially processed and where memory is limited in treatment capacity.

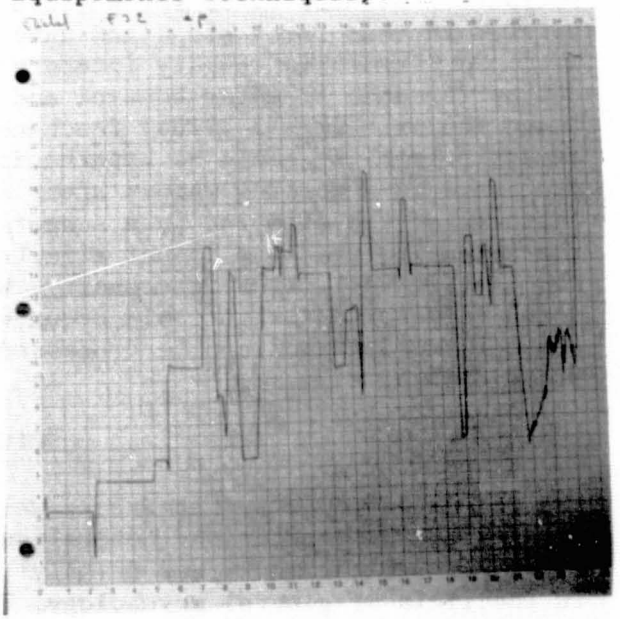
Finally, the electrotactile stimulation could be used more as an alerting signal than a complex information feedback. It does not seem to be well adapted to code semantic or mental image information. Visual or auditory information, however, does benefit from the transfer function in the peripheral structure where information could be treated before its access to a working memory. The visual dominance effect could be the result of the advantage of both a transfer function and a sensory memory register where information is pretreated and memorized for a short time. Dividing attention has an effect on the acquisition of the information but not on the subsequent decision processes, suggesting that in motor tasks the attention variable is primarily affecting early cognitive activities.

ACKNOWLEDGMENTS

I wish to thank H. H. Kwee and N. Quétin for their help and comments during the preparation of this paper. This research was supported by the DRET (Direction de la Recherche des Equipements Techniques).



Recording of the summed forces with the visual feedback.



Recording of the summed forces with the electro-tactile feedback.

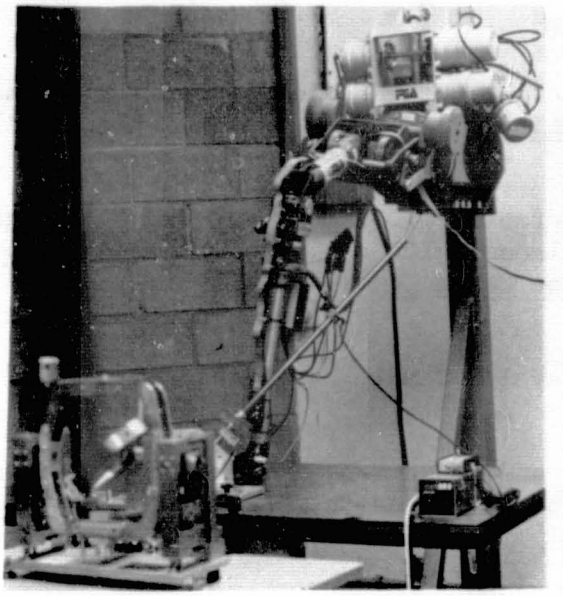


Photo 1

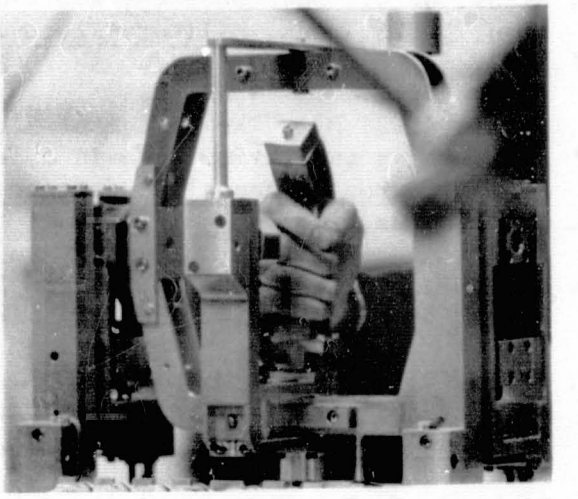


Photo 2

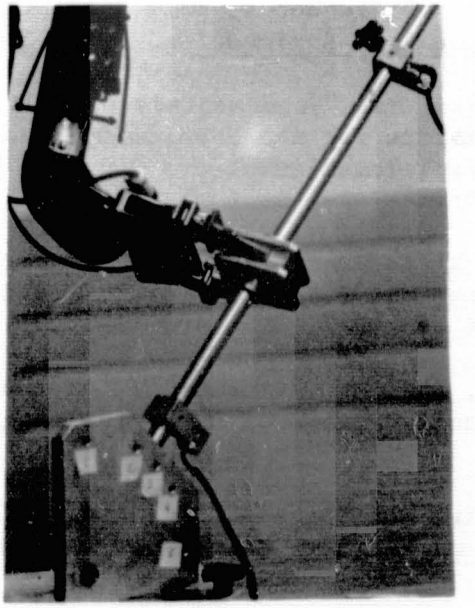


Photo 3

REFERENCES

1. Bliss, J. C., and Kotovsly, K., "Tactile Presentation of Visual Information," IEEE Trans. Military Electronics, Vol. MIL-7, pp. 108-115, April, 1963.

2. Gaillard, J. P., "The effect of Electrotactile and Visual Force Feedback on Learning and Performance in a Telemanipulation Task," Proceedings of the 1st European Annual Conference on Human Control and Decision, Delf. Netherlands, 1981.

3. Hill, J. W., "Sutdy to Design and Develop Remote Manipulator Systems," Report 5 and 6, S.R.I., 1977.

4. Kato, I., Kumamoto, M., Tamura, S., and Tsunekawa, Y., "Human Cognitional Ability for Electrical Stimulation Signals," Advances in External Control of Human Extremities, Belgrade, Yugloslavia, 1970a: 68-84.

5. Richard, T. F., "L'Attention", PUF, 1980.

6. Shannon, G. F., "Sensory Feedback," Progress Report No. 14, 42-53, Bio-Engineering Institute of New-Brunswick, Fredericton, Canada, 1976.

7. Solomonow, M., Lyman J., and Freedy, A., "Electrotactile Two-Point Discrimination as a Function of Frequency, Body Site, Laterality and Stimulation Codes," Annals of Biomedical Engineering, 1977, 5 (1): 47-60.

8. Solomonow, M. and Lyman, J., "Fundamentals of Multichannel Sensory Feedback from Telemanipulators for the Physically Handicapped," Conference Internationale Sur Les Telemanipulateurs Pour Les Handicapes Physiques. IRIA, Rocquencourt, France, 4-6 September, 1978.

9. Szeto, A. Y., Prior, R. E., and Lyman, J., "Electrocutaneous Pulse Rate and Pulse with Psychometric Functions," ISPO World Congress on Prosthetics and Orthoties, May 26 - June 2, New York, 1977.

10. Tachi, S., Tanie, K., Abe, M., "Experiments on the Magnitude Sensation of Electrocutaneous Stimuli," Bull. of Mechanical Engineering Laboratory n° 30 (1978).

11. Treisman, A. M., Davies A., "Divided Attention to Ear and Eye," in S. Kornblum, Attention and Performance, New York Academic Press, 1973, IV, 101-117.

D13

MAN/MACHINE INTERFACE DEVELOPMENT FOR THE REMOTEX CONCEPT*

J. Garin
Fuel Recycle Division
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37830

M. M. Clarke
Oak Ridge Associated Universities
Oak Ridge, Tennessee 37830

SUMMARY

This paper describes ongoing research at the Oak Ridge National Laboratory (ORNL) to develop a man/machine interface system that can be used to remotely control a system composed of a transporter base and a force-reflecting, servo-controlled manipulator. A unique feature of the concept is the incorporation of totally remote operation. Thus, a major objective is the requirement that an operator have a sense of presence in the remote environment. Man/machine interface requirements for this totally remote operation remain to be developed. Therefore, a simulator is being built to optimize such requirements. These developments are the subject of this paper.

INTRODUCTION

Work on the conceptual design of a pilot-scale plant for reprocessing breeder reactor fuels is underway at ORNL. The plant design will meet all current federal regulations for reprocessing plants and will serve as a prototype for future production plants. A unique feature of this concept is the incorporation of totally remote operation and maintenance of the process equipment within a large barn-like hot cell. This approach, called REMOTEX, uses servomanipulators coupled with television viewing to extend human capabilities into the hostile cell environment. The REMOTEX concept provides significant improvements for fuel reprocessing plants and other nuclear facilities in the areas of safeguarding nuclear materials, reducing radiation exposure to humans, improving plant availability, recovering from unplanned events, and plant decommissioning.

*Research sponsored by the Office of Nuclear Fuel Cycle, U.S. Department of Energy, under Contract No. W-7405-eng-26 with Union Carbide Corporation.

The success of the REMOTEX concept is probably most dependent on the performance of the remote manipulation, viewing, and maintenance system. A review of other facilities indicates that using more versatile manipulation will result in decreased complexity (and cost) of in-cell equipment and the downtime associated with maintenance. This indication can be substantiated in a number of ways. For example, increasing the dexterity of remote maintenance manipulators to approach that of direct handling may reduce the cost of designing and building such equipment.

Numerous studies and tests have defined dexterity as the time required to perform sets of work tasks using different types of remote manipulators. The general results of these studies are presented in Ref. 1. Dexterity depends on force reflection (kinesthetic sensing), tactile feedback, visual feedback, and audio feedback. Therefore, we hypothesized that by maximizing properly filtered feedback, the effectiveness of our teleoperated systems would also be maximized. Optimally, the man/machine interface provides information feedback from the work environment that allows the operator to react in "real time" and to provide feedforward information into the work place.

APPROACH TO SIMULATOR DEVELOPMENT

The development of the man/machine interface at ORNL began with a state-of-the-art literature search of studies on related systems. This search revealed a need for emphasizing visual systems and the types of teleoperated servomanipulator controls.

Conceptual design of the ORNL system will continue with the testing of prototypes. The prototypes are evaluated at the subsystem level described in Ref. 2 and then undergo a full system test in a simulated environment (the simulator).

The simulator is an integrated man/machine interface that contains all of the subsystems in a modular form. These subsystems consist of imaging closed circuit television, graphic display, audio system, manipulator controls, camera controls and selectors, transporter controller, and auxiliary control systems.

This man/machine interface is then coupled to a remote handling system and a simulated nuclear fuel recycling environment. A series of generic tasks have been selected using a detailed time and motion study based on similar facilities and the current conceptual design.

The evaluation and selection of each man/machine subsystem will then depend on system test evaluations and the performance of the teleoperated system.

STATUS

The man/machine interface simulator is 90% complete (see Figs. 1 and 2). Figure 1 shows an artist's concept of the optimized system using one operator man/machine interface. Figure 2 shows the man/machine interface module as currently assembled.

FUTURE PLANS

Immediate plans for future work include the following:

1. Design of in-cell, small-volume work tasks appropriate to a simulated chemical process module;
2. Evaluation of small-volume coverage cameras for large-volume usage; and
3. Evaluation of small- and large-volume cameras for obstacle avoidance.

Longer-range goals include system optimization for the following:

1. Larger volume, less dextrous manipulator tasks; and
2. Transporter deployment.

REFERENCES

1. J. Garin, "Remotex and Servomanipulator Needs in Nuclear Fuel Reprocessing Plants," *Proc., of Workshop to Delineate the Economic, Technical, and Policy Issues for the Remote Maintenance in Energy Systems*, University of Florida, Gainesville, Florida, March 1981.

2. M. M. Clarke and J. Garin, "Identification of Cognitive Factors Related to Remote Work Performance Using Closed Circuit TV Displays," *Proc. 17th Annual Conference on Manual Control*, Los Angeles, June 16-17, 1981.

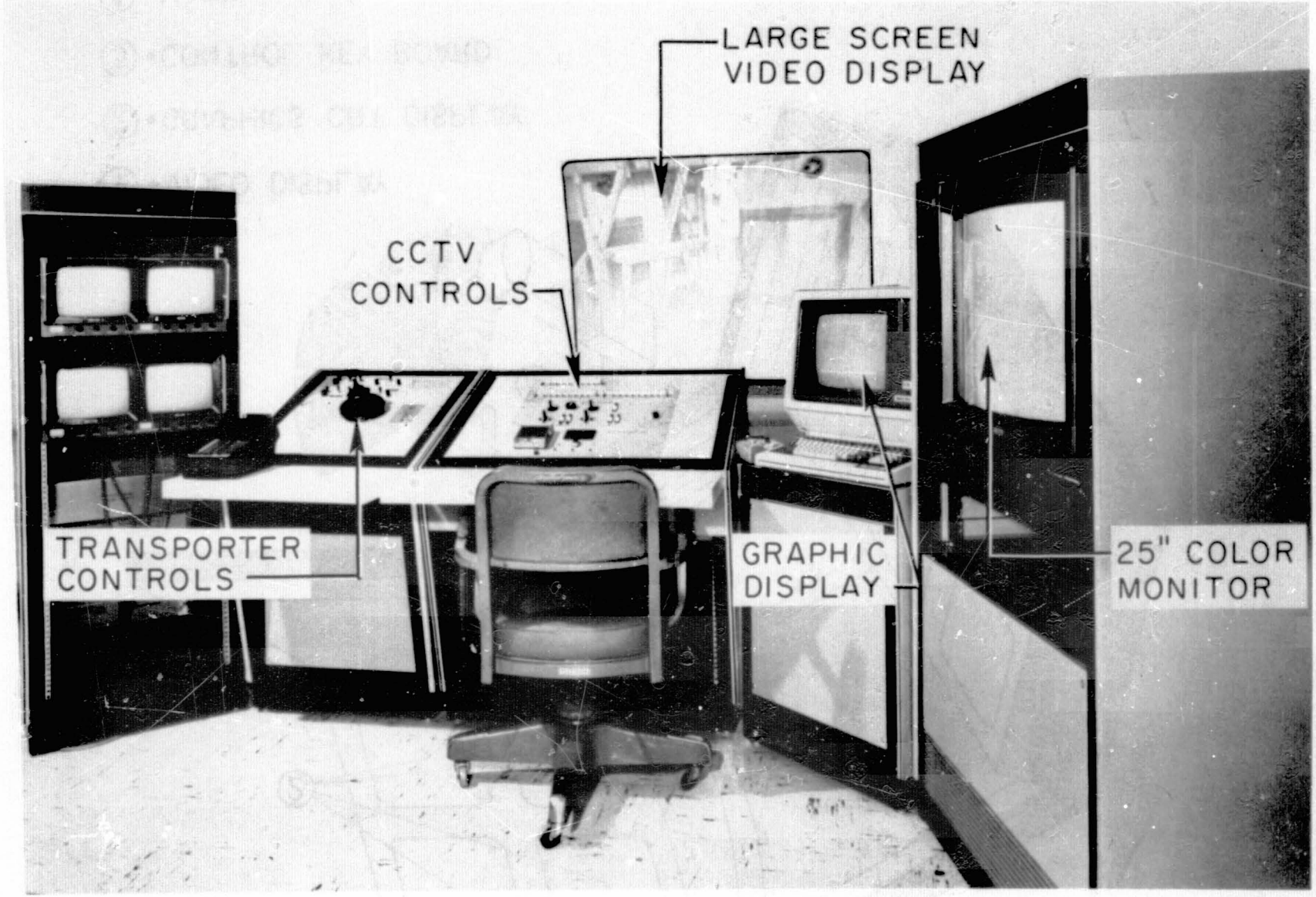
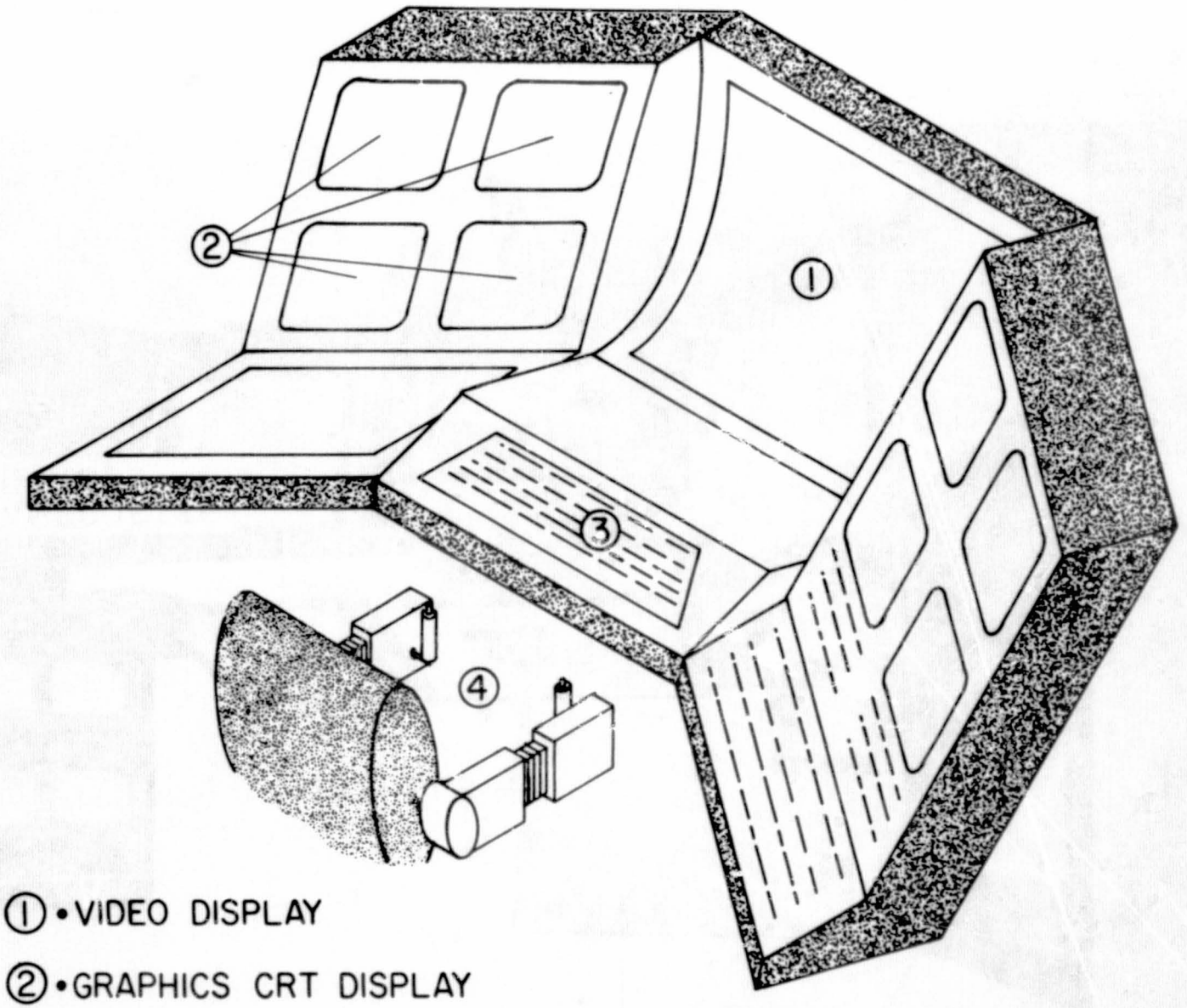


Fig. 1

ORNL DWG.80-15031



① • VIDEO DISPLAY

② • GRAPHICS CRT DISPLAY

③ • CONTROL KEY BOARD

④ • MANIPULATOR CONTROLLER

Fig. 2

D14

IDENTIFICATION OF COGNITIVE FACTORS RELATED TO REMOTE WORK PERFORMANCE USING CLOSED CIRCUIT TV DISPLAYS*

M. M. Clarke
Oak Ridge Associated Universities
Oak Ridge, Tennessee 37830

J. Garin
Fuel Recycle Division
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37830

SUMMARY

This paper describes research at the Oak Ridge National Laboratory which identifies operator perceptual-cognitive styles as predictors of remote task performance. Of particular interest are those remote tasks which require the use of servo-controlled master-slave manipulators and closed circuit television for teleoperator repair and maintenance of nuclear fuel recycling systems. At the present time data analysis is incomplete; therefore, this paper suggests a useful procedure for identifying such perceptual styles, rather than giving definitive results.

INTRODUCTION

The particular aspects of man/machine interface for the ORNL fuel recycle program have been described previously (1). This facility will employ remote viewing and manipulator control systems. In general, however, manipulator control using closed circuit television, involving the remote movement of mass with limited visual information, is analogous to other remote movement tasks such as vehicle piloting. In both situations the operator or pilot moves mass at a distance under some constraint in regards to available visual information. Therefore, it seemed reasonable to assume that previous research concerning perceptual-cognitive abilities as related to piloting/navigation tasks would be applicable to the present problem, particularly since there is a dearth of literature concerning perceptual cognitive abilities relating to manipulator control skills, Generally, people who can function perceptually in the presence of misleading, distracting, or conflicting information have been found to be superior at various remote control tasks: navy pilots (2), nonpilot navigators and intercept operators (3), and automobile drivers (4,5).

---

*Research sponsored by the Office of Nuclear Fuel Cycle, U.S. Department of Energy, under contract No. W-7405-ENG-26 with Union Carbide Corporation.

EXPERIMENTAL PROCEDURES

To evaluate the relationship between perceptual-cognitive functioning and remote handling ability, a group of inexperienced subjects were administered a battery of perceptual-cognitive tests to measure perceptual functioning in the presence of conflicting or distracting perceptual inputs. Subjects were then briefed and received preliminary training in the use of manipulators for remote handling tasks using closed circuit television displays as the primary mode of information feedback. It was hypothesized that subjects' perceptual test scores would predict training results regarding level of performance and speed of learning.

Subjects

Forty-seven employees (26 male and 21 female) of Oak Ridge National Laboratory or Oak Ridge Associated Universities volunteered as subjects. No subject had prior manipulator experience. All subjects received an in-depth vision screening for acuity and phorias (near and far) and depth and color perception. Subjects were asked to fill out a biographical questionnaire which contained many items with possible relevance to remote task performance. Items covered age, sex, preferred hand (right or left), occupation, education, sports activities, and television viewing habits. A question was also included concerning motion sickness history since previous literature (3) has indicated that such sickness may result from the way subjects react to conflicting perceptual inputs and may be predictive of remote control training results.

Groups Perceptual-Cognitive Tests

All subjects were administered three group perceptual tests prior to exposure to remote tasks:

1. Hidden Figures Test, which measures flexibility of closure and which is related to the "field independence" cognitive style (6).

2. Paper Folding Test, which measures the ability to mentally restructure a visualized figure into components, and involves the serial mental rotation of a spatial configuration in short-term memory (6).

3. Manikin Test, an experimental group perceptual test refined at Oak Ridge Associated Universities, which demands that subjects make rapid decisions about the spatial orientation of visual stimuli.

Remote Test Station Layout

Subjects performed the experimental remote tasks with master-slave M-8 manipulators. A 25-in. TV monitor was placed between the master arms at eye level on a solid partition simulating a division between the task (cell) area and the operating area. The operating area was diffusely lighted and surrounded by canvas curtains. A chair was placed at a distance of four times the monitor height from the monitor. To keep viewing distance constant, the subjects were asked to remain seated while they used the manipulators. The layout was measured in S.I. units.

The slave side of the partition was also surrounded by canvas curtains and was brightly lighted at approximately 343 candelas/meter (100 foot-lamberts). Intermittent measures on illumination were taken throughout the testing. The board on which the individual tasks were performed was painted gray-blue.

Television cameras were installed in positions 45° above horizontal in the sagittal plane and at an offset position at 30° elevation and 45° to the left side of the sagittal plane.

Test Procedures for Remote Tasks

Subjects were given a demonstration in the use of a manipulator and then allowed to practice freely for 30 min. placing twelve, 2.5-cm wooden cubes in a 30-cm high steel pot. They were then asked to perform fairly simple remote tasks, three trials for each.

Task 1. Build a 3-2-1 pyramid with the cubes.

Task 2. Take out and replace 7.5-cm high wooden geometrical shapes from a form board.

Task 3. Unscrew, place on table, pick up, and rescrew a 2.5-cm bolt from a steel plate.

All tasks were performed with the camera at the offset position. Subjects were then asked to repeat the bolt test with the camera at 45° above horizontal, which unpublished work by the first author has shown to be a disorienting view.

Subjects returned individually, approximately four weeks later, and repeated the total remote task sequence.

ANALYSIS AND RESULTS

The predictive validity of the group perceptual tests, and of the biographical items is being assessed in terms of the results of remote task performance. Preliminary regression analysis indicates that the Hidden Figures Test is a useful predictor, accounting for a significant proportion of the variance in measures of remote training outcome. The relationship between this test and training outcomes remains when any confounding effect of sex is removed from the data. (This procedure is necessary because women often perform more poorly on such tests.) Results, therefore, suggest the usefulness of group perceptual cognitive tests such as Hidden Figures, in future remote operator selection and training programs.

REFERENCES

1. Garin, John, and Clarke, Margaret M.; "Man/Machine Interface Development for the Remotex Concept," Proc. 17th Annu. Conf. on Manual Control, Los Angeles, June 1981.

2. Long, G.M.; "Rod and Frame Test Performance Among Naval Aviation Personnel," Percept. Mot. Skills, 41: 950, (1975).

3. Kennedy, Robert S.; "Motion Sickness Questionnaire and Field Independence Scores as Predictors of Success in Naval Aviation Training," Aviation Space and Environmental Medicine, 1975, 1349-52.

4. Barrett, G.V. and Thornton, C.C.; "Relation Between Perceptual Style and Driver Reaction to an Emergency Situation," J. Appl. Psychol., 52: 169-76, (1968).

5. Harano, R.M.; "Relationship of field dependence and motor vehicle accident involvement," Percept. Mot. Skills, 31: 272-74, (1970).

6. Ekstrom, R.B., French, John W., and Harmon, H.H.; Manual for Kit of Factor References Cognitive Tests, Educational Testing Service, Princeton, New Jersey, 1976.

D15

SIMPLE GEOMETRIC ALGORITHMS TO AID IN CLEARANCE MANAGEMENT FOR ROBOTIC MECHANISMS

E. L. Copeland
L. D. Ray
J. D. Peticolas

Lockheed Engineering and Management Services Company, Inc.

ABSTRACT

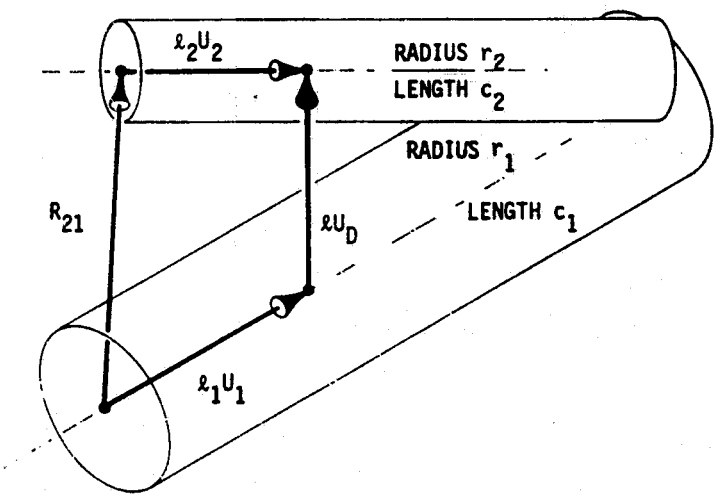
For Robotic Mechanisms which are required to operate in quarters limited by external structures, the problem of clearance is often of considerable interest. In such cases it is possible to distinguish between "contact prediction" and "minimum safe clearance management."

The advantage of the distinction is principally in the fact that the latter may be quite simple and well suited to real-time calculation, whereas the former may require more precision, sophistication, and computational overhead.

This paper deals with the selection of global geometric shapes such as lines, planes, circles, spheres, cylinders, etc., and the associated computational algorithms which provide relatively inexpensive estimates of minimum spatial clearance for safe operations. The Space Shuttle, Remote Manipulator System, and the Power Extension Package are used as an example.

SIMPLE GEOMETRIC ALGORITHMS TO AID IN CLEARANCE MANAGEMENT FOR ROBOTIC MECHANISMS

CONTACT PREDICTION (HIGH RESOLUTION)	MINIMUM SAFE CLEARANCE MANAGEMENT
o REQUIRES ACCURATE, DETAILED GEOMETRIC MODELING	o ALLOWS THE USE OF GLOBAL GEOMETRIC SHAPES FOR THE ENCLOSURE OF LARGE PIECES OF STRUCTURE
o REQUIRES ACCURATE MOTION MODELING	o ALLOWS MORE TOLERANCE IN MOTION MODELING
o COMPARATIVELY HIGH COMPUTATIONAL TIME AND LARGE PROGRAM SIZE	o COMPARATIVELY LOW COMPUTATIONAL TIME AND SMALL PROGRAM SIZE
o USUALLY INVOLVES BASIC GEOMETRIES (POINTS, SPHERES, CUBES)	o USUALLY INVOLVES MORE COMPLEX GEOMETRIES (CIRCLES, CYLINDERS, CONES, SPATIAL TRAJECTORIES, SURFACES OF REVOLUTION)



$$U_D = \text{UNIT}(U_1 \times U_2)$$

$$R_{21} + \ell_2 U_2 = \ell_1 U_1 + \ell U_D$$

$$U_1 \cdot U_D = 0$$

$$U_2 \cdot U_D = 0$$

TO DETERMINE MINIMUM CLEARANCE

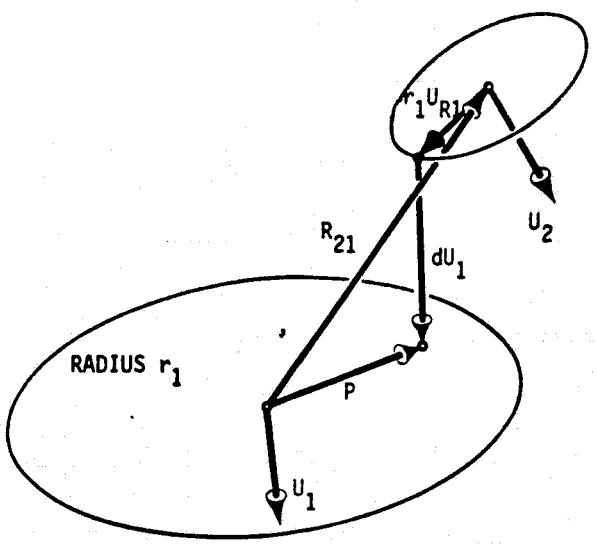
$$d = \ell - r_1 - r_2$$

$$d > 0$$

$$0 \leq \ell_1 \leq c_1$$

$$0 \leq \ell_2 \leq c_2$$

FIGURE 1.- CYLINDRICAL-SHELL-TO-CYLINDRICAL-SHELL CLEARANCE.



$$U_{R1} = U_2 \times (\text{UNIT}(U_1 \times U_2))$$

$$R_{21} + r_1 U_{R1} + dU_1 = P$$

$$U_1 \cdot P = 0$$

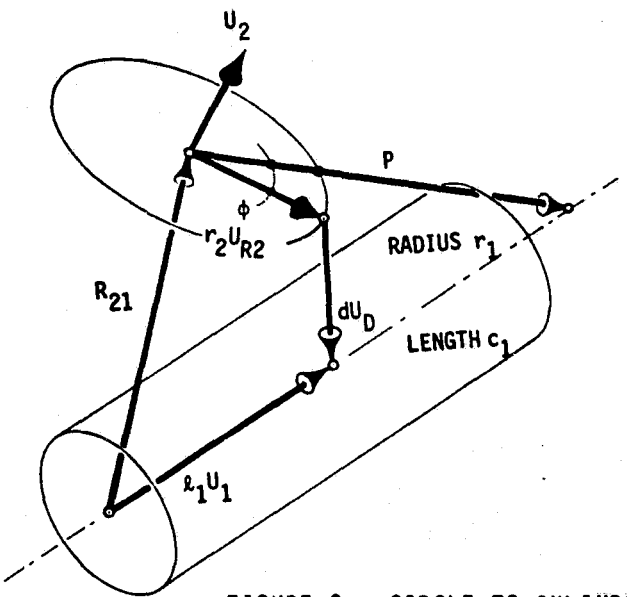
$$p = \text{LENGTH}(P)$$

TO DETERMINE MINIMUM CLEARANCE

$$d > 0$$

$$p < r_1$$

FIGURE 2.- CIRCLE-TO-PLANE CIRCULAR AREA CLEARANCE.



$$R_{21} + r_2 U_{R2} + dU_D = \ell_1 U_1$$

$$U_D \cdot \frac{d}{d\phi} U_{R2} = 0$$

$$U_D \cdot U_1 = 0$$

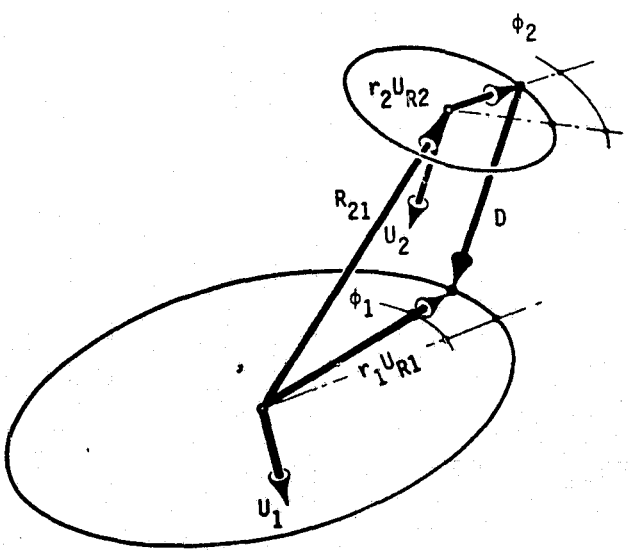
ADJUST ϕ TO MINIMIZE d

TO DETERMINE MINIMUM CLEARANCE

$$d > r_1$$

$$0 < \ell_1 < c_1$$

FIGURE 3.- CIRCLE-TO-CYLINDRICAL-SHELL CLEARANCE.



$$R_{21} + r_2 U_{R2} + D = r_1 U_{R1}$$

$$D \cdot \frac{d}{d\phi_1} U_{R1} = 0$$

$$D \cdot \frac{d}{d\phi_2} U_{R2} = 0$$

D = LENGTH (D)

ADJUST ϕ_1, ϕ_2 TO MINIMIZE d

TO DETERMINE MINIMUM CLEARANCE

$$D \cdot U_1 > 0$$

FIGURE 4.- CIRCLE-TO-CIRCLE CLEARANCE.

CLEARANCE MANAGEMENT MODEL APPLICATION TO SHUTTLE/ REMOTE MANIPULATOR/POWER EXTENSION PACKAGE (PEP)

- PEP CONTROL SYSTEM EXECUTES TWO-AXIS MOTION SWEEPING OUT DISK-LIKE CYLINDER
- RMS SECTIONS CYLINDRICAL
- SHUTTLE SECTIONED AND MODELED AS CYLINDERS AND LINES
- THE ALGORITHMS AND AN EXECUTIVE LOGIC WERE USED TO AID IN DESIGNING OPERATIONAL TRAJECTORIES AND CONFIGURATIONS FOR THE POWER EXTENSION PACKAGE AS WELL AS OTHER SHUTTLE PAYLOADS

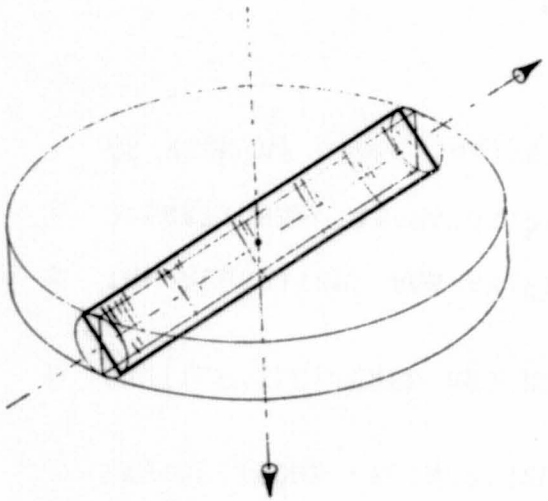


FIGURE 5.- CYLINDRICAL VOLUME SWEPT BY POWER EXTENSION PACKAGE DEPLOYED.

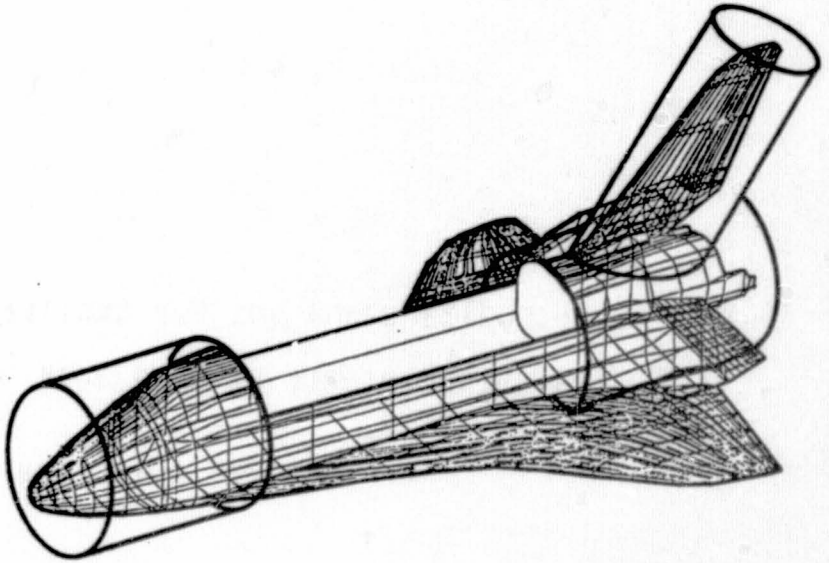


FIGURE 6.- CLEARANCE MANAGEMENT MODEL CYLINDRICAL ENCLOSURE OF SHUTTLE ORBITER.

CONCLUSIONS AND EXTENSIONS

- THE DISTINCTION BETWEEN "CONTACT PREDICTION" AND "MINIMUM SAFE CLEARANCE MANAGEMENT" LED TO A SIGNIFICANT SAVING IN COMPUTATIONAL TIME AND COMPLEXITY FOR THIS APPLICATION
- THE FOUR SIMPLE GEOMETRIC COMBINATIONS DISCUSSED HEREIN HAVE BEEN INTEGRATED INTO AN ALGORITHM TO MANAGE THE OVERALL CLEARANCE BETWEEN CYLINDERS
- THE CIRCLE AND CYLINDER DEMONSTRATION HEREIN MAY BE GENERALIZED TO SURFACES OF REVOLUTION AND CONES, ETC., FOR MORE COMPLICATED APPLICATIONS
- EVALUATION AND SELECTION OF SUITABLE DISPLAYS AND CUES FOR MANUAL CONTROL SHOULD BE INVESTIGATED

THE INFLUENCE OF MOTION CUES ON DRIVER-VEHICLE PERFORMANCE IN A SIMULATOR

By Brian S. Repa and Philip M. Leucht
Engineering Mechanics Department
General Motors Research Laboratories

and

Walter W. Wierwille
Department of Industrial Engineering and Operations Research
Virginia Polytechnic Institute and State University

SUMMARY

Four different motion base configurations, ranging from no motion to a configuration consisting of roll plus yaw plus attenuated lateral translation, were studied on the Virginia Polytechnic Institute (VPI) driving simulator. The same trio of differently responding vehicles was simulated on each motion configuration; and the effects of the vehicle characteristics on driver-vehicle system performance, driver control activity, and driver opinion ratings of vehicle performance during driving were compared for the different motion configurations. Analysis of the data showed that:

1. The effects of changes in vehicle characteristics on the different objective and subjective measures of driver-vehicle performance were not disguised by the lack of physical motion.
2. Studies performed on a fixed-base simulator can thus be used to draw inferences despite the lack of motion.
3. The presence of motion tended to reduce path keeping errors and driver control activity.
4. If limited physical motions were to be added to a driving simulator, roll and yaw are recommended because of their marked influence on driver-vehicle performance and their relative ease of implementation.
5. The importance of motion increased as the driving maneuvers became more extreme.

INTRODUCTION

This research was undertaken to determine what influence motion cues have on the sensitivity of driver performance measures to changes in vehicle response characteristics. The findings were expected to be useful to researchers considering the inclusion of motion in their simulator facilities, in interpreting the results of research conducted in fixed-base facilities, and in improving our understanding of the cues that drivers use in controlling their cars. Most of the available information about the

value of motion cues comes from research performed in aircraft simulators. A notable exception to this is the work performed by McLane and Wierwille (reference 1) which did show that driver performance is augmented by the addition of motion cues. However, they did not address the question of whether or not the same relative trends are observed with different motion systems but only that absolute measures of performance do differ.

Repa et al (reference 2) compared the performance of three research engineers in a fixed-base simulator and a Variable Response Vehicle under different vehicle response conditions and found very close correspondence between simulator and full-scale testing. However, it can be argued that these particular drivers were more skilled than the average driver and were able to adapt to the different vehicle configurations and to the laboratory simulator more effectively and more quickly than average drivers. Therefore, differences between the performance of the different vehicle configurations as well as between the full-scale and the simulator results may have been minimized because of the select group of drivers that was used.

The evidence currently available from aircraft simulator studies generally favors the presence of motion. According to Rolfe (reference 3), "... this conclusion is based upon findings such as:

(a) Experienced pilots perform better in a moving simulator than they do in a static simulator.
(b) The pattern of the pilot's contol activity in an aircraft is more closely approximated in a moving rather than a static simulator.
(c) Good motion systems are seen as increasing the 'realism' and acceptability of the simulation."

Additional points regarding the importance of motion are made by Staples (reference 4):

"It is important where an unstable or near neutrally stable vehicle has to be controlled (e.g., a bicycle).

It assists in control at high frequency, particularly where precision or high accuracy is desired (e.g., in tracking).

It may be required because of its effect on visual acuity and visual judgment.

It may degrade performance because of coupling either with a human control mode (pilot-induced oscillations), even on a vehicle with stable open-loop dynamics, or with the natural frequency of body components (vibratory responses).

Finally, because motion provides direct information about the nature of the response to the pilot and because it cannot be ignored (in the short term), it serves as a powerful alerting input, directing attention to or eliciting an instinctive reaction to the response even though it was not the center of attention at that time."

In spite of the more or less general acceptance of the above statements, there is still considerable controversy concerning just how much motion is required under what vehicle conditions and for which tasks

(reference 5). Unfortunately, there is no simple basic framework of human behavior which will allow us to identify in any given situation the elements of the total environment which contribute to the pattern of behavior (reference 4). As a result, the following research was undertaken to provide information of direct relevance to the cues drivers use during normal highway driving.

METHOD

Driving Simulator

The facility that was used for this research was designed and built at Virginia Polytechnic Institute and State University under a General Motors grant. The simulator provides the driver with an unprogrammed* television-type display in coordination with the motions of yaw, roll, pitch, and lateral and longitudinal translation. Four channels of sound along with vibration are also provided. A brief description of the visual and motion systems follows. More complete descriptions are given in references 6 and 7.

Visual System - Generation of the simulated roadway image (figure 1) is accomplished by a special purpose computing system containing 37 integrated circuits. The generated signals are initially displayed on the face of a cathode ray tube. A TV camera scan converts this image and transmits it by cable to a TV monitor mounted above and behind the dash on the simulator buck. A Fresnel lens with an effective focal length of 50.8 cm (20.0 in), located between the monitor and the human operator, decreases the apparent roadway image proximity to the driver and enhances the illusion of distance [effective distance 10.1 m (33 ft)]. Additional realism is provided by a plexiglass windshield and a sheet metal mock-up representing a hood and fenders immediately in front of the dash. The field of view provided by the TV monitor and lens subtends 39° vertically and 48° horizontally. During simulation, all room lights are turned off so that only the roadway display and the illuminated speedometer are visible to the driver.

Motion System - The simulator is composed of an upper and lower platform, three main struts, and four motion servos (figure 2). The upper platform consists of a standard automotive configuration including bucket seat, dashboard (with speedometer), steering wheel, brake and accelerator pedal, and the visual display equipment described above. The upper platform is pivoted at each end which permits roll motion about an axis 33.0 cm (13.0 in) from the upper platform floor. The roll motion is accomplished by the roll servo which is attached between the upper and lower platforms. The lower platform, while providing support for the upper platform, is supported by nine precision rubber-wheeled casters which enable the platforms to rotate in yaw and to translate longitudinally and laterally on a

---

* The visual display is a result of the actions of the driver and is not a pre-programmed scenario.

large lucite sheet. The platform motions are generated by the combined action of three servo-operated struts which have one end pivoted about a floor anchored support. Two of the servos are aligned with the longitudinal axis of the simulator and provide longitudinal translation and rotation in yaw. The third servo is oriented laterally and provides the lateral translation. Each motion servo is monitored by a feedback potentiometer and is controlled by its own electromechanical valve which receives signal inputs from the analog computer. A hydraulic pump, regulated at 10.4 MPa (1500 psi), provides fluid power to the motion system. Acoustical insulation of the pump unit controls the noise level in the simulator room. A 2.5 gallon accumulator and associated valving are used to maintain constant fluid pressure at any required flowrate. Fluid temperature is controlled by an integral refrigeration-type heat exchanger.

Vehicle Simulation - Simulation of vehicle dynamics is performed by an EAI TR-48 analog computer. Driver inputs to the steering wheel, accelerator, and brake pedals are sensed by potentiometers and converted to electrical signals. These signals are the inputs to the vehicle model simulated on the computer. Outputs of the model are analog voltages of vehicle velocity, roll, yaw, lateral position, and longitudinal acceleration, which form the signals applied to the motion servos. These signals are also applied to the driver's speedometer and to the image generation circuitry which continuously adjusts the visual display characteristics (position, perspective, velocity, etc.) to correspond to the simulated vehicle state.

The vehicle model used for the simulation allowed for rotations in roll, yaw, and pitch, as well as lateral and longitudinal translations. Three separate inputs were provided; namely, steering wheel displacement, accelerator/brake displacement, and aerodynamic force (wind gust). The model consisted of a set of transfer functions relating the three inputs to the vehicle motion components. This approach approximates the dynamic response of the vehicle and permits matching of the model responses to either measured full-scale responses or to responses generated by digital simulation codes.

Experimental Design

The primary concern of this research was whether or not the addition or removal of motion cues would change the performance trends that are observed over different vehicle response configurations. The capability of the VPI simulator includes four separate motion cues -- roll, yaw, lateral translation, and longitudinal translation -- and any of their combinations. At the outset, it was decided to focus only on the lateral motion information so the longitudinal cue was removed for all the tests. In the interest of experimental design simplicity, meaningfulness, and time, a total of four motion base configurations and three vehicle response configurations were chosen for study, resulting in twelve experimental conditions as shown in Table 1. While other combinations of motion cues would have been possible, this particular set of configurations (fixed-base, roll only, roll plus yaw, and roll plus yaw plus attenuated lateral translation)

represents a natural progression in complexity and, from a hardware development point of view, a way to develop a motion platform over a period of time.

The "base" vehicle response configuration is that of a typical compact car and the "slow" and "fast" configurations were obtained by selecting tires with cornering stiffnesses that might be expected to lie at the extremes of the after market tire range. Slight adjustments in the cornering stiffnesses were made in order to keep the vehicle understeer and, hence, the control sensitivity* constant for all three vehicle configurations. As an added convenience, the aerodynamic center of pressure was assumed to be located at the front wheels, which is a much easier configuration to represent with the transfer function model of the vehicle used on the VPI simulator. The corresponding vehicle responses to control and disturbance inputs are shown in figure 3.

In the interest of obtaining unbiased driver performance in the various motion base configurations, the experiment was designed to provide each subject with only one motion base configuration. Thus, four different groups of six subjects were randomly assigned to each of the four motion configurations (for a total of twenty-four subjects). Half of the subjects in each group were female.

The schedule of events during the course of an experimental run is shown in figure 4. The first experimental run was used for practice.

Performance Measures

Random Wind Gus Disturbance Task - RMS measures of lateral position, yaw angle, and steering wheel angle were obtained on-line using hybrid circuitry for the last 2.5 minutes of the random wind gust task. The number of steering reversals greater than 3.4 deg were also obtained online over the same time period. RMS lateral position and yaw angle are direct measures of driver-vehicle lateral control performance. RMS steering wheel angle and steering reversals are interpreted as measures of the driver's steering control effort.

Step Wind Gust Disturbance Task - Maximum lane position deviation was used as the performance measure in this task. The maximum deviation was measured from the actual vehicle position prior to gust onset and not from the center of the lane.

Subjective Ratings - A 0-10 rating scale was used by the subjects at the completion of both the random and step wind gust disturbances to rate the performance of each vehicle with respect to the disturbances that were encountered. ("0" corresponded to no noticeable disturbance while "10" corresponded to an uncontrollable disturbance.)

---

* A vehicle with greater control sensitivity, which is defined as the lateral acceleration produced by a steering wheel angle of 100 degrees, produces the same maneuver with a smaller steering wheel input.

RESULTS

Subjective Ratings

A highly significant difference in subjective ratings ($p < 0.0001$) occurred as a result of changes in vehicle response characteristics. No significant differences were noted as a function of motion base configuration. Figure 5 shows the mean ratings as a function of vehicle characteristics.

Random Wind Gust Disturbance Performance

Highly significant differences in performance for all four objective measures occurred as a result of changes in vehicle response characteristics ($p < 0.001$). No significant differences were observed as a function of motion base configuration for lateral position deviations and steering reversals. Figures 6 and 7 show the influence of vehicle configuration on these two measures.

For steering wheel deviations, a significant effect ($p < 0.02$) was also noted as a function of motion base configuration. Figure 8 shows the effects of both vehicle characteristics and motion base configuration on steering deviations.

For yaw angle deviations, significant effects due to motion configuration as well as to the interaction between motion configuration and vehicle characteristics were observed ($p < 0.01$). These effects are shown in figure 9.

Step Wind Gust Disturbance Performance

Significant effects due to vehicle configuration ($p < 0.0001$) and to motion base configurations ($p < 0.03$) were noted for the maximum lane position measure. Figure 10 shows these effects graphically.

DISCUSSION

We do not, in a strict sense, know which of the responses to changes in vehicle characteristics in the various motion cue configurations are most accurate because we have not compared the results with real driving. It is assumed, however, that the responses in the roll plus yaw plus attenuated lateral translation configuration come closest to those in the real driving environment.

While it is clear that some measures of driver performance in a simulator are significantly influenced by the combination of motion cues that are presented, it is equally clear that meaningful results can still be obtained with a fixed base facility. Several important aspects of driver-vehicle performance -- namely, subjective opinion ratings, RMS lane deviations, and steering reversals -- were not affected by the different motion configurations. Others, i.e., steering wheel deviations for the

random disturbance task and maximum lane deviations for the step disturbance task, were affected by motion; but the relative ranking of the different responding vehicles remained the same for all motion configurations. Only yaw angle deviations for the random disturbance task showed some divergence between the different vehicle configurations as motion cues were eliminated. For this measure, "slow" responding vehicles had disproportionately higher deviations in the no motion configuration. This would be a far more serious concern if the absence of motion tended to disguise rather than exaggerate an effect that was observed when motion was present. Previous studies have shown that yaw angle deviations are one of the most sensitive measures to changes in experimental conditions (reference 7), so the present results are not surprising. Also, aircraft simulator studies have shown that motion cues become more important when the task becomes more demanding (reference 4) - which corresponds, in this case, to the control of the "slow" responding vehicles. Table 2 summarizes the high degree of transferability of results that exists across the different motion configurations.

The reductions in lane tracking deviations and driver control activity that did occur when motion was added can be attributed to the alerting effect provided by the motion. This effect is most clearly illustrated by the marked reduction in maximum lane deviations that took place when yaw motion was added in the step disturbance regulation task (Fig. 10). The same effect is well known in the field of aircraft simulation where the presence of motion in a wind gust regulation task leads not only to earlier recognition and reaction to the gust, thus reducing the initial excursion, but also results in more accurate control of the motion (reference 8). In the absence of motion, the human operator must obtain the information on the vehicle's response to the disturbance primarily from the visual display. Some alerting effect is possible from auditory cues, but the direction for the required correction must still come from the visual display.

The transient response characteristics of the vehicle which were varied to provide the different responding vehicles for the present study are subtle vehicle characteristics to examine on a fixed base simulator. If control sensitivity were the characteristic being varied, it would automatically produce changes in driver performance whether or not motion cues were present. As a result, the present study was expected to be particularly effective in revealing the limitations of a fixed base facility. The fact that the fixed-base configuration provided results close to those of the moving base configuration indicates that at least for normal highway driving meaneuvers, motion cues may provide information that is in many ways redundant to that provided by visual cues.

The lack of statistically significant differences in summary performance measures with such a small subject pool (six subjects per motion configuration) must be viewed with some caution. Had more subjects been used or a more detailed analysis of the data been performed, more of the differences may have become significant. The attenuated lateral motion is also best thought of as only a partial cue because of the attenuation that is required to keep lateral excursions within a limited physical range [i.e., ± .5 m (20 in)]. If a larger throw had been used, it is possible that significant changes would have been observed with the addition of

lateral motion. However, even with nonsignificant differences in performance across motion configurations, there is no denying that the presence of motion does result in a more realistic driving simulation.

The marked reduction in maximum lane excursions following a step wind gust that occurred when yaw motion was introduced does suggest that as the driving maneuvers become more severe, certain motion cues may become more important. For example, in a skid control maneuver, the absence of a yaw motion cue might very well lead to a delayed driver response and subsequent loss of control with a vehicle configuration that would otherwise be well within the control capabilities of the driver. The high degree of transferability of results from a fixed base configuration to moving base configurations that was obtained in the present study thus applies to normal highway driving maneuvers only.

REFERENCES

1. McLane, R. C. and Wierwille, W. W., "The Influence of Motion and Audio Cues on Driver Performance in an Automobile Simulator," Human Factors, 17(5), 1975.

2. Repa, B. S., Alexandridis, A. A., Howell, L. J., and Wierwille, W. W.," The Influence of Vehicle Control Dynamics on driver-Vehicle Performance," Proc. 5th VSD - 2nd IUTAM Symposium on the Dynamics of Vehicles on Roads and Tracks, Vienna, pp. 320-333, Swets Zeitlinger, Amsterdam, 1977.

3. Rolf, J. M., "Keeping Up on the Ground," Aeronautical Journal, July 1977.

4. Staples, K. F., "Current Problems of Flight Simulators for Research," Aeronautical Journal, January 1978.

5. Gundry, J., "The Effectiveness and Sophistication of Motion Cues Provided in Flight Simulators," Human Operators and Simulation, The Chameleon Press, Ltd., London, 1977.

6. Wierwille, W. W., "Driving Simulator Design for Realistic Handling," Proceedings of the Third International Conference on Vehicle System Dynamics, Swets and Zeitlinger, B. V., Amsterdam, 1975.

7. Repa, B. S., and Wierwille, W. W., "Driver Performance in Controlling a Driving Simulator with Varying Vehicle Response Characteristics," SAE Transactions, Volume 85, pp. 2453-2468, 1976.

8. Perry, D. H. and Naish, J. M., "Flight Simulation for Research," The Journal of the Royal Aeronautical Society, October 1964.

TABLE 1
EXPERIMENTAL CONDITIONS

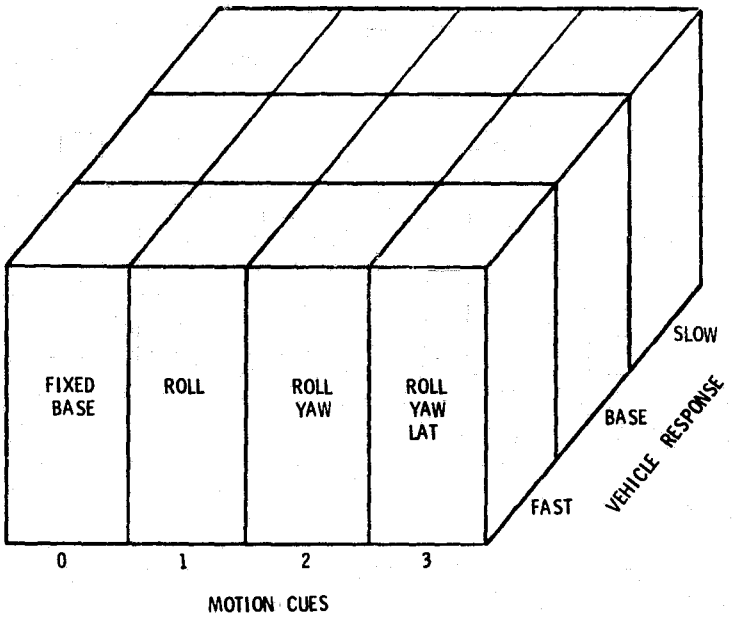


TABLE 2
TRANSFERABILITY OF RESULTS SUMMARY

Dependent Variable	Transferability of Results Across Motion Configurations
Subjective Ratings	Transferable on an Absolute Basis
Lane Deviations (Random Gusts)	Transferable on an Absolute Basis
Steering Reversals (Random Gusts)	Transferable on an Absolute Basis
Maximum Lane Deviations (Step Gusts)	Transferable on a Relative Basis
Steering Wheel Deviations (Random Gusts)	Transferable on a Relative Basis
Yaw Angle Deviations (Random Gusts)	Not Directly Transferable

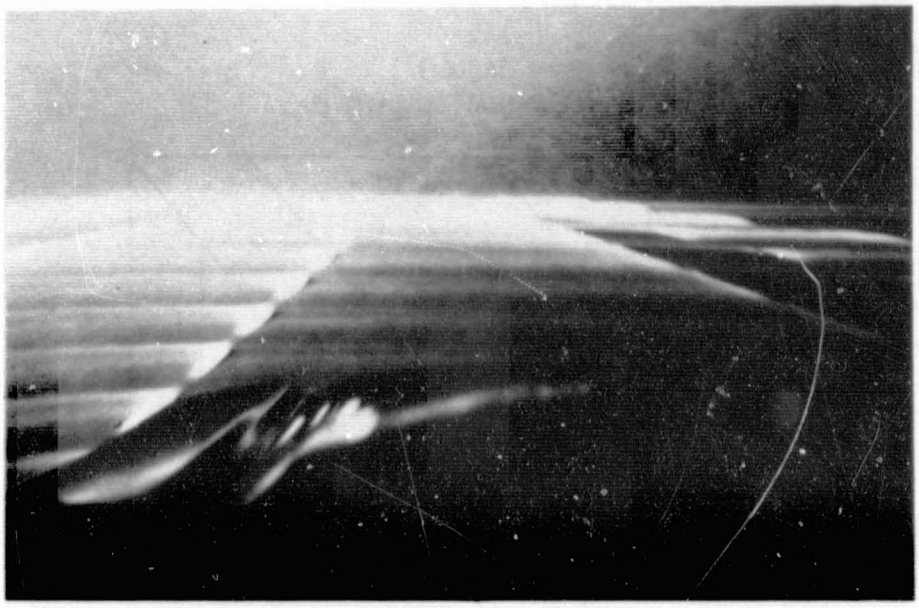


FIG. 1 PHOTOGRAPH OF ROADWAY IMAGE

FIG. 2 DRIVING SIMULATOR MOTION PLATFORM

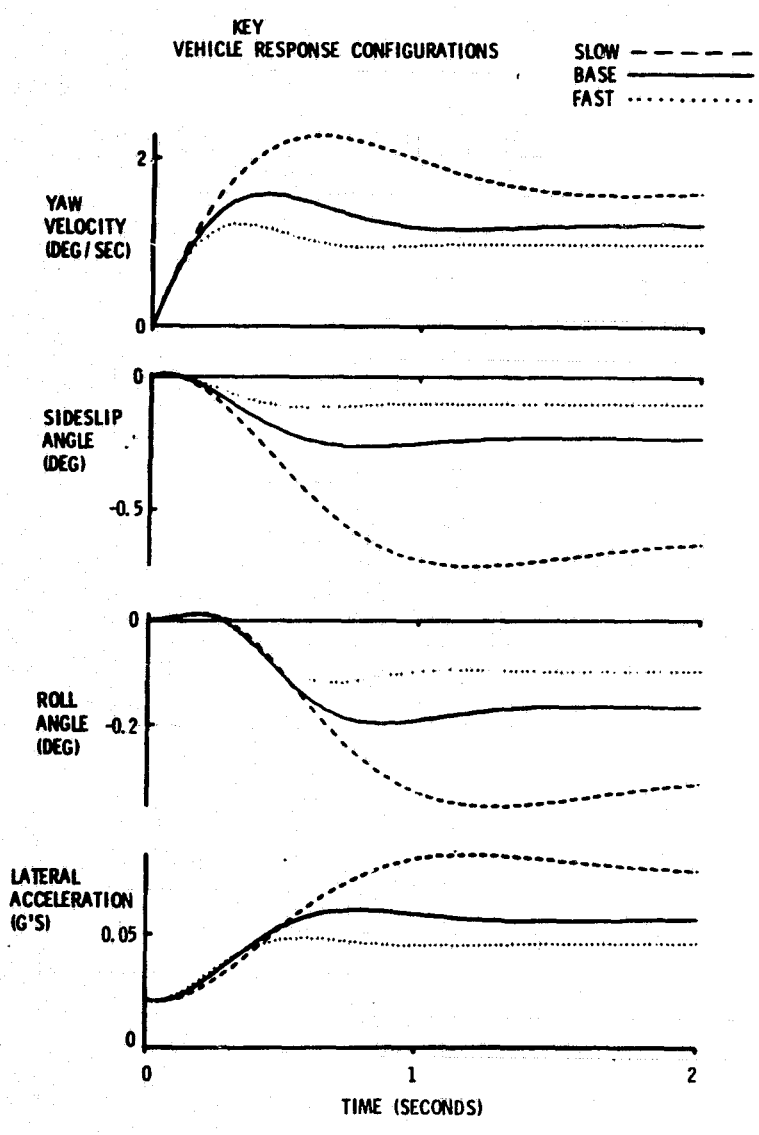


FIG. 3B VEHICLE RESPONSES TO DISTURBANCE INPUTS

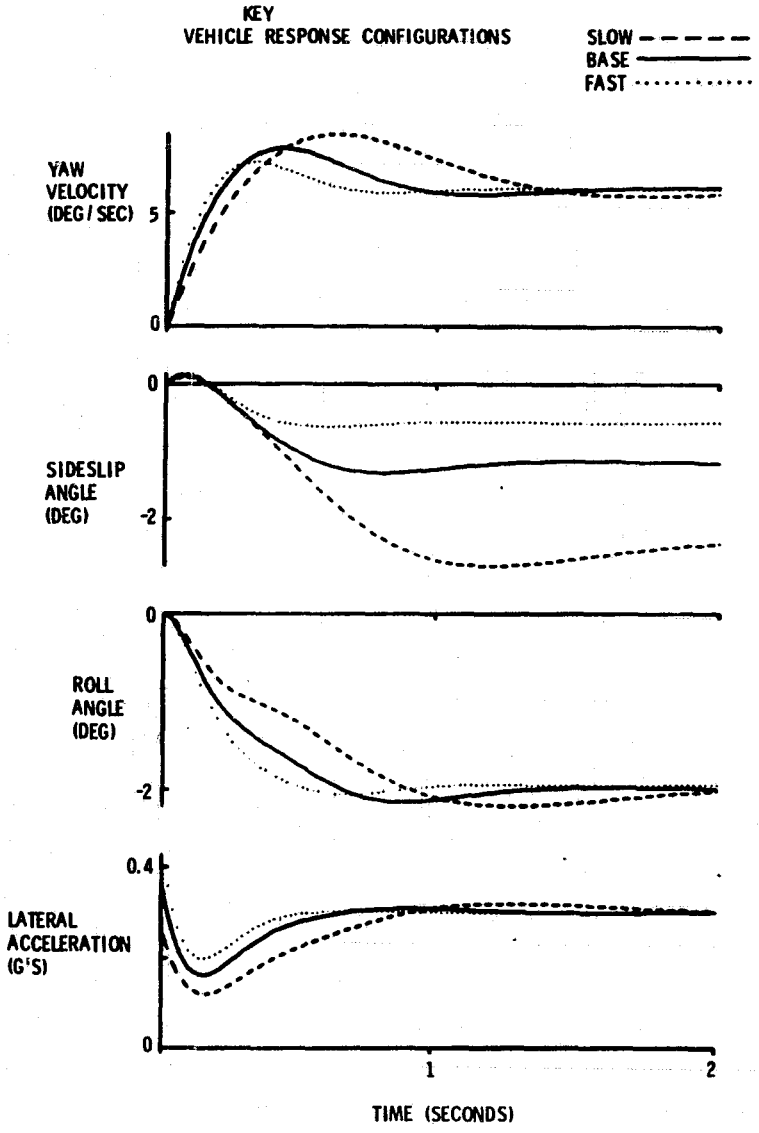


FIG. 3A VEHICLE RESPONSES TO CONTROL INPUTS

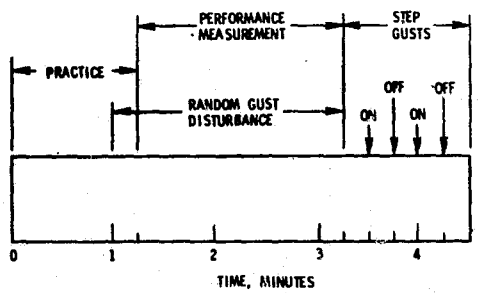


FIG. 4 SCHEDULE OF EVENTS

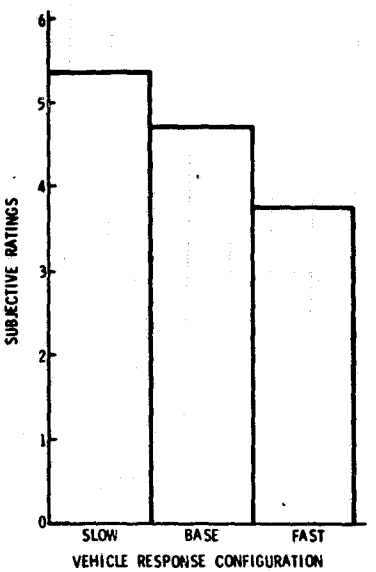


FIG. 5 SUBJECTIVE RATINGS FOR THE THREE VEHICLE RESPONSE CONFIGURATIONS

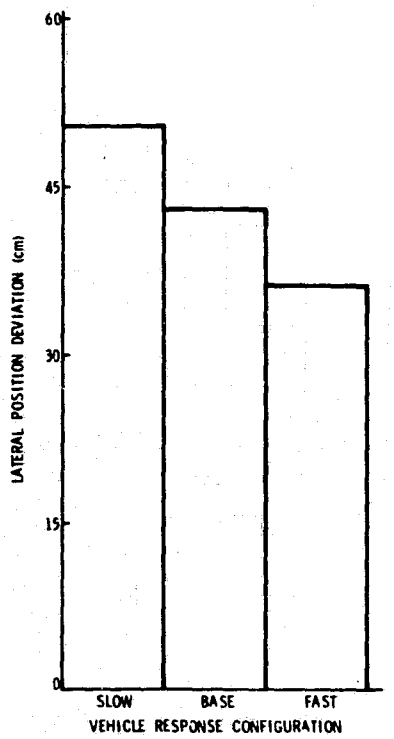


FIG. 6 LATERAL DEVIATIONS DURING RANDOM DISTURBANCES FOR THREE VEHICLE RESPONSE CONFIGURATIONS

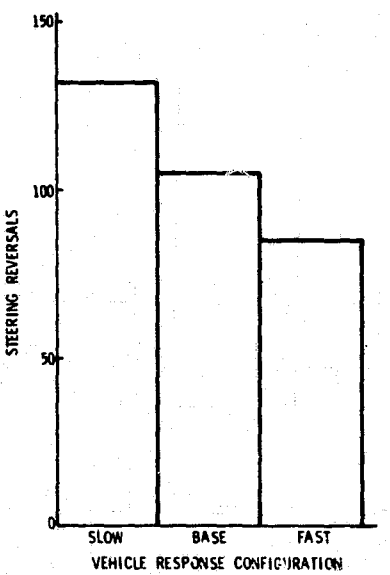


FIG. 7 STEERING REVERSALS DURING RANDOM DISTURBANCES FOR THREE VEHICLE RESPONSE CONFIGURATIONS

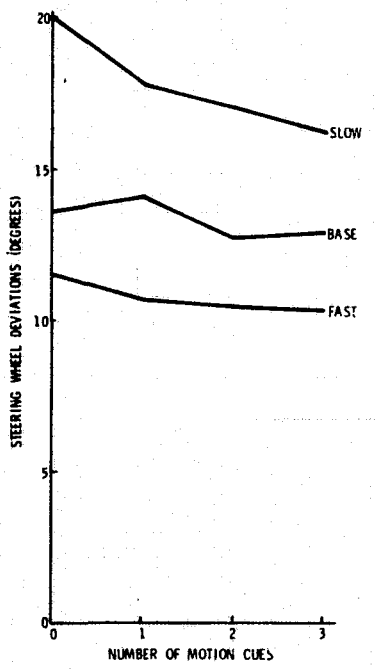


FIG. 8 STEERING DEVIATIONS DURING RANDOM DISTURBANCES FOR FOUR MOTION CUE CONDITIONS AND THREE VEHICLE RESPONSE CONFIGURATIONS

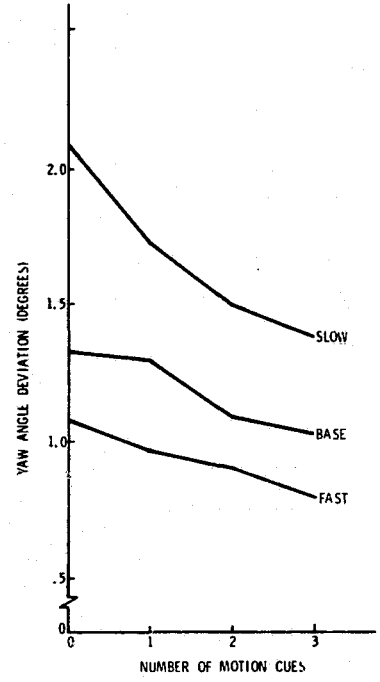


FIG. 9 YAW DEVIATIONS DURING RANDOM DISTURBANCES FOR FOUR MOTION CUE CONDITIONS AND THREE VEHICLE RESPONSE CONFIGURATIONS

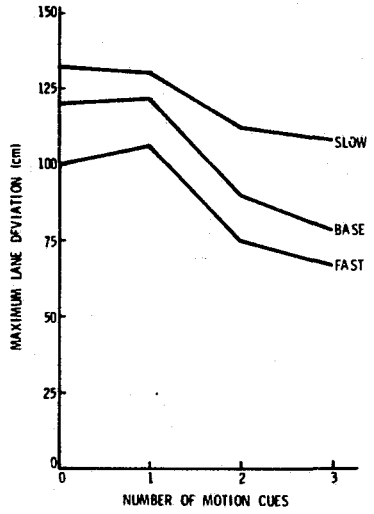


FIG. 10 MAXIMUM LANE DEVIATIONS DURING STEP DISTURBANCES FOR FOUR MOTION CUE CONDITIONS AND THREE VEHICLE RESPONSE CONFIGURATIONS

MULTIPLE MAN-MACHINE INTERFACES

By Lillian Stanton, Ph.D. and Creighton W. Cook

Lockheed-California Company

SUMMARY

Modern technology has produced machines that, in many instances, can see, hear, and touch with greater accuracy and precision than human beings. Consequently, the military pilot is more a systems manager, often doing battle against a target he never sees. This paper explores the multiple man-machine interfaces inherent in military pilot training, their social implications, and the issue of possible negative feedback.

INTRODUCTION

A man-machine system is a combination of one or more human beings and one or more physical components interacting to bring about, from given inputs, some desired output. Man interacts with the system to fulfill the function for which the system is designed.

Ever since prehistoric man discovered the wheel and other "necessary inventions," man-machine interfaces have proliferated. Nowhere today is this more evident than in the multiple machine interfaces with which the military student pilot must contend. Modern technology has produced machines that, in many instances, can see, hear, measure, and touch with greater accuracy and precision than human beings. Consequently, the military pilot is more a systems manager, often doing battle against a target he never sees.

Although many of the same interfaces are faced by flight crews and multi-engine pilots, this paper will address only the multiple machine interfaces encountered by the undergraduate jet pilot trainee, a one on one situation.

Current plans are underway to upgrade military training aircraft in order to facilitate the student's transition to the highly sophisticated tactical aircraft. Not only must the student contend with complex display symbology, formatting, and data rate, but increasingly more of his training takes place in devices that are, in some ways, more sophisticated than the aircraft itself. His academic training is effected largely through computer-assisted instruction, often combined with video disc technology. This system is linked to a computerized training management system which combines management and scheduling functions with information storage and retrieval. In addition, both the aircraft and the training devices are equipped with performance measurement equipment which records the student's every movement.

Today, we joke about the scarf, goggles, and needle-ball indicator era, but that has a core of truth. In the early days of flying, when the flying machine was relatively simple (figure 1), the man-machine interface was straightforward. Superchargers and oxygen systems were non-existent. Rarely out of sight of land, the pilot used railroads, rivers, and other distinctive landmarks as navigational guides. His instruments were simple - - perhaps a compass, altimeter, needle-ball indicator for horizontal reference, air speed indicator, fuel gauge, and oil pressure gauge (figure 2). Dead reckoning or "seat-of-the pants" flying was the norm. There were no radio communications. The pilot was in control (figure 3).

Three-quarters of a century later, aviation has changed drastically. As man's knowledge of aerodynamics, aircraft structures, and electronics has increased, airplanes have developed accordingly. Because modern aircraft are complex, with many separate systems to be monitored, their crowded cockpits contain numerous readouts. Because panel space is often insufficient, space-saving devices such as two or three pointer instruments have been introduced. However, often these innovations are at the cost of reading difficulty and therefore may cause increased errors.

Today, cathode ray tube displays (CRT) permit the display, as on a television screen, of only the information which is currently required. Alternatively, basic flight information may be superimposed on the real world by a head-up display (HUD) which projects the CRT image on a screen in front of the pilot. The HUD permits the pilot to maintain his view of the real world and still receive vital information without having to look down at the instrument panel. It has an additional advantage of superimposing

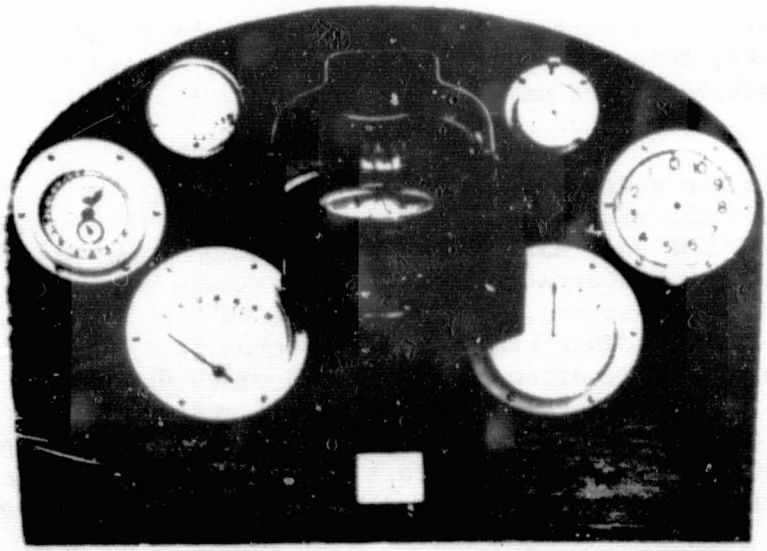


Figure 1. Early Cockpit - Very Simple

C-3

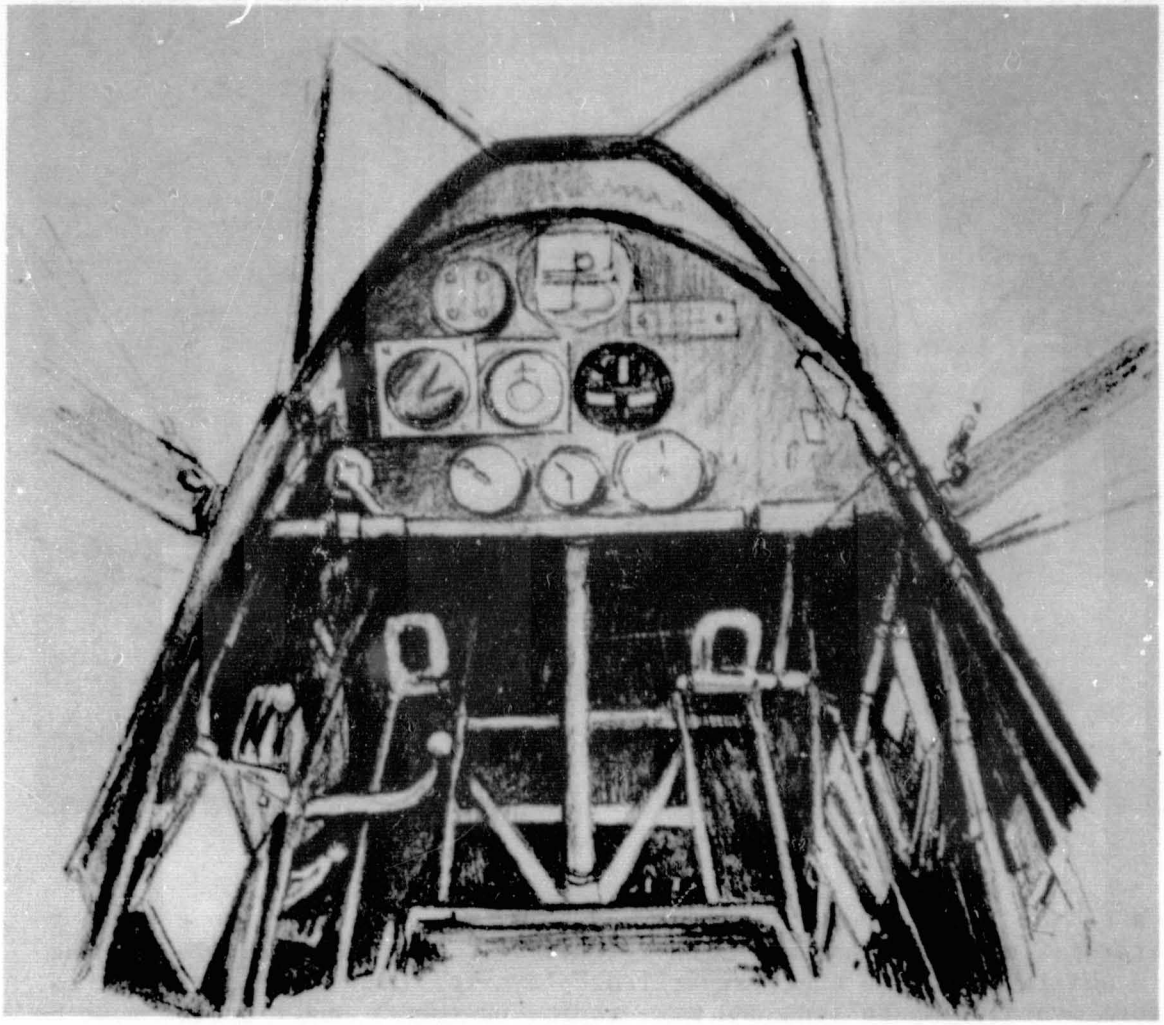


Figure 2. Sketch of Early Instrument Panel

the real world on the synthetic picture collimated to infinity so that the pilot does not have to change his focus. However, an inadequate field of view of the real world results in HUD clutter.

A helmet-mounted sight provides a simple display in front of the operator's eye, and is designed to improve his ability to aim weapons or sensors at potential targets by merely looking in the target's direction. Sensors on the helmet feed data to a computer which calculates where the pilot is looking. The helmet-mounted display is a more complex device which produces an instrument display in front of one of the pilot's eyes.

Despite the growing awareness that there is a tendency to put too much information on the display, an airplane such as the F/A-18 weapons system

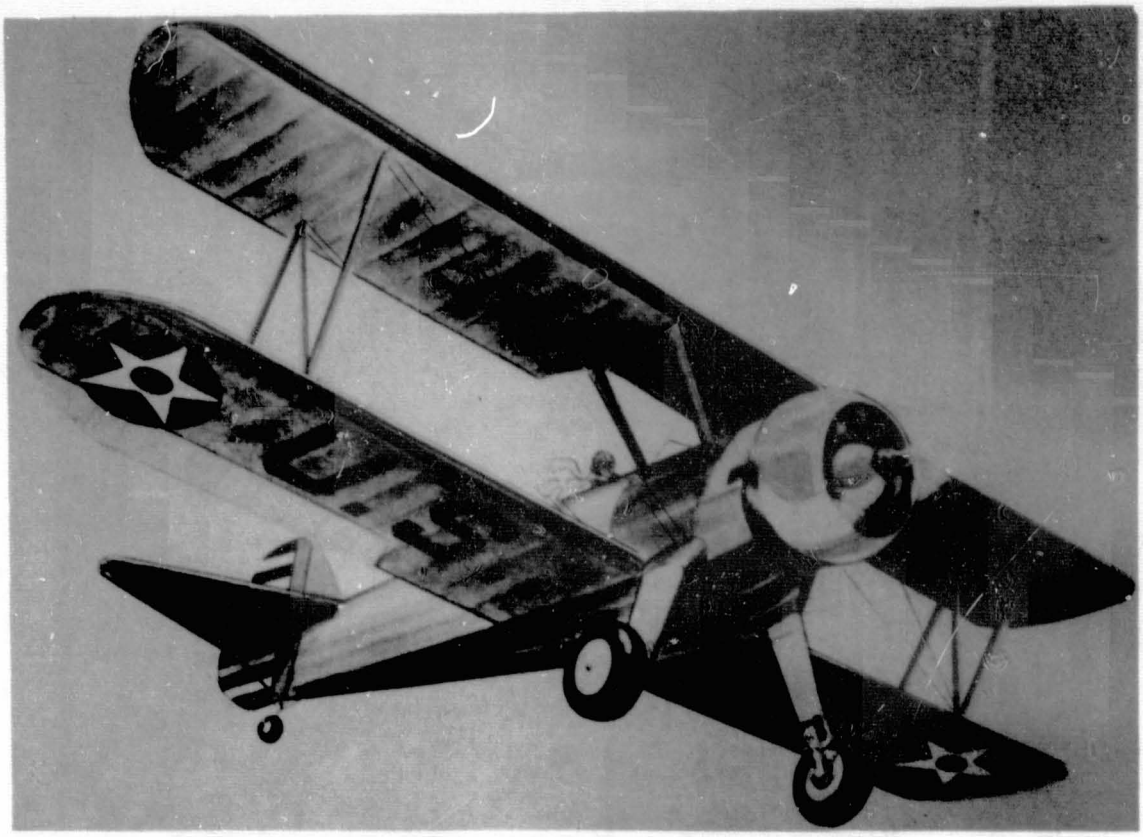


Figure 3. Early Trainer

has a capability of over 700 display symbols which the pilot must learn to interpret, integrate, and act upon (figure 4). Cognitive factors such as man's ability to cross-check several displays for inconsistencies, and to perform mental computations and make rapid judgements under stress must be considered. There are questions as to whether all the displayed information is necessary to a pilot, whether it is correct, or how much of it can process. It is not always necessary to know the precise state of the system, rather whether the system is within limits. And, even when the pilot must know the precise value of the parameters he is monitoring, he may need to check it only once or twice during the flight. Although the human information processing system (figure 5) is analogous to an electrical communications system, man is essentially a single channel device with a processing rate of approximately two bits per second. It may be pointless to present him simultaneously with all the information he may need throughout an entire flight when he can process only a relatively small amount of that information at a time.

Ongoing work in training technology includes a refinement of speech synthesis, speech recognition and speech digitizers along with increased capacity artificial intelligence and computer-aided decision making. The relative merits of such new technology -- the value of speech recognition,

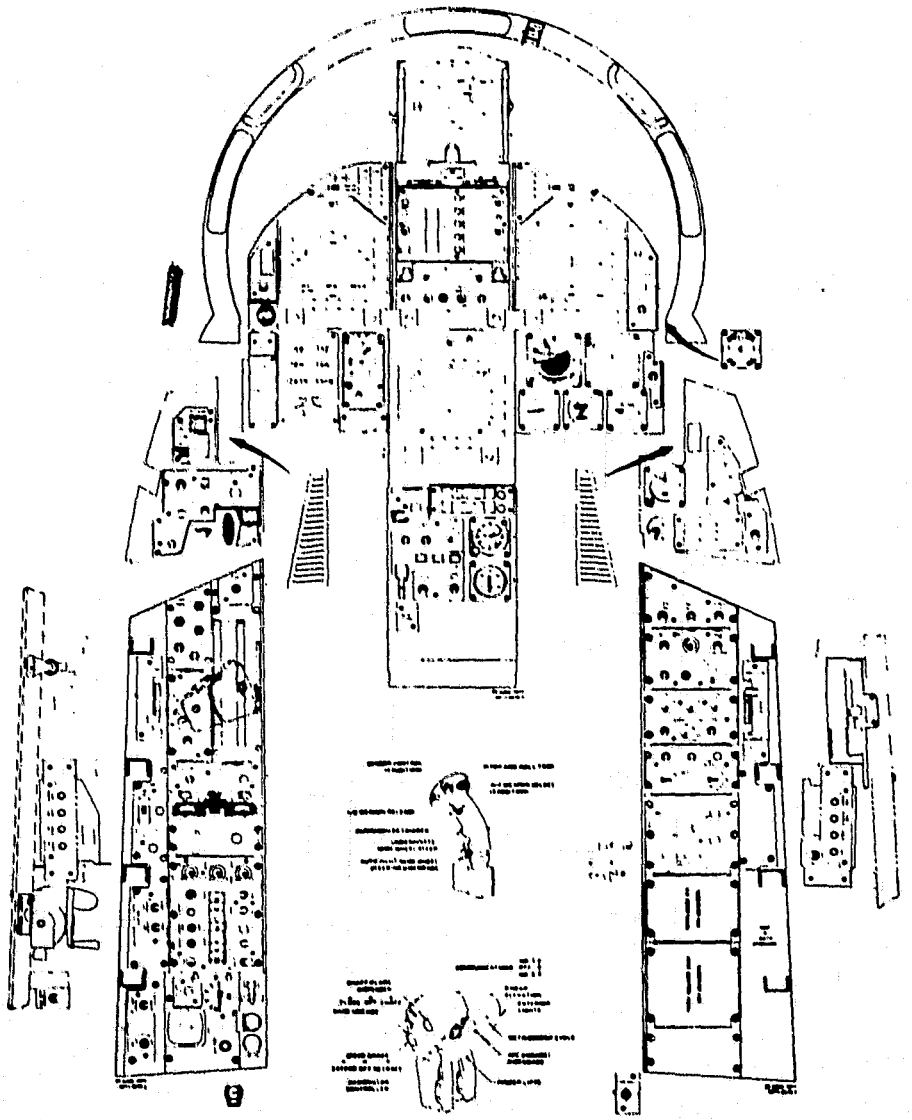


Figure 4. F-18 Crew Station

the merits of CRTs with side switches versus touch panels for controls, the value of pictorial formats such as pathways in the sky, and the feasibility of combining traffic information, weather radar, and a map display on the electronic horizontal situation indicator without information overload -- have yet to be determined (figure 6).

Some experts predict that the future pilot will be presented with more and more information which will be utilized less and less. And, as the cost of displays escalates and the pilot's work load increases, the overall efficiency of this man-machine system will decline due to the shortcomings of many of the new display devices with regard to man's capacities (figure 7).

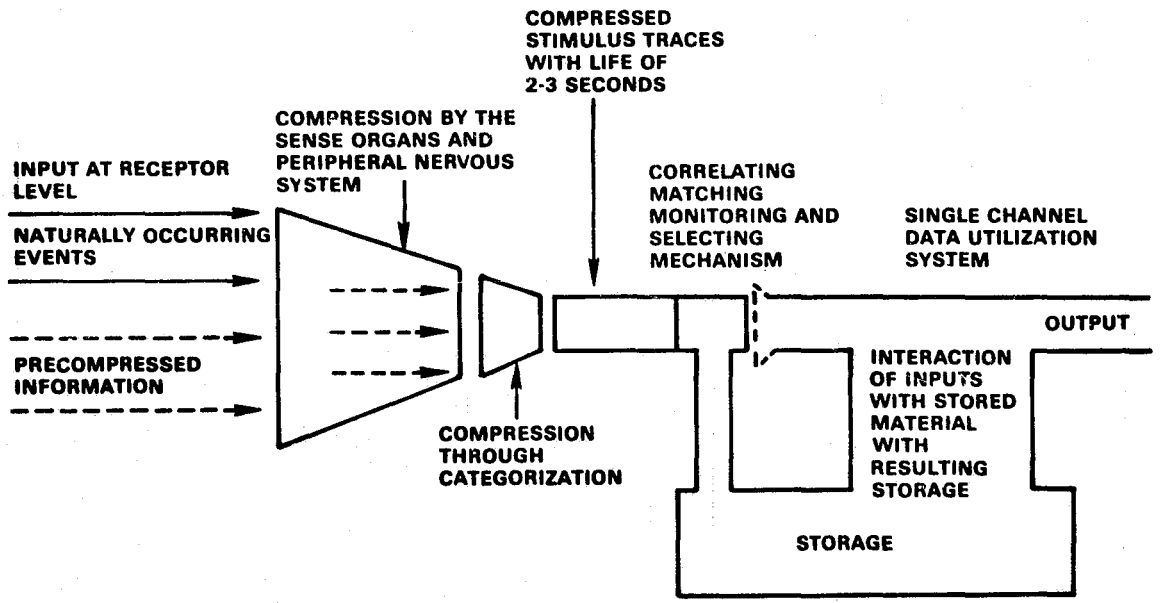


Figure 5. Information Processing

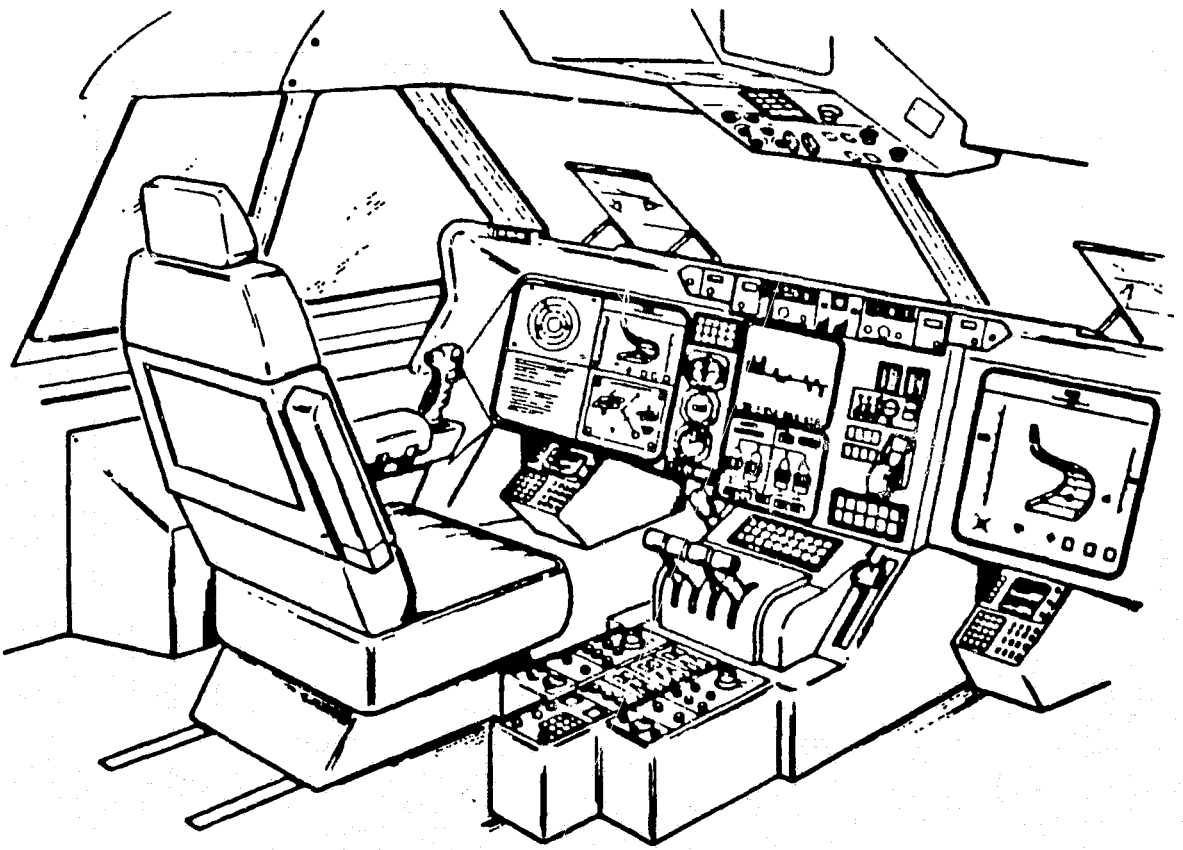


Figure 6. An Advanced Concepts Flight Station

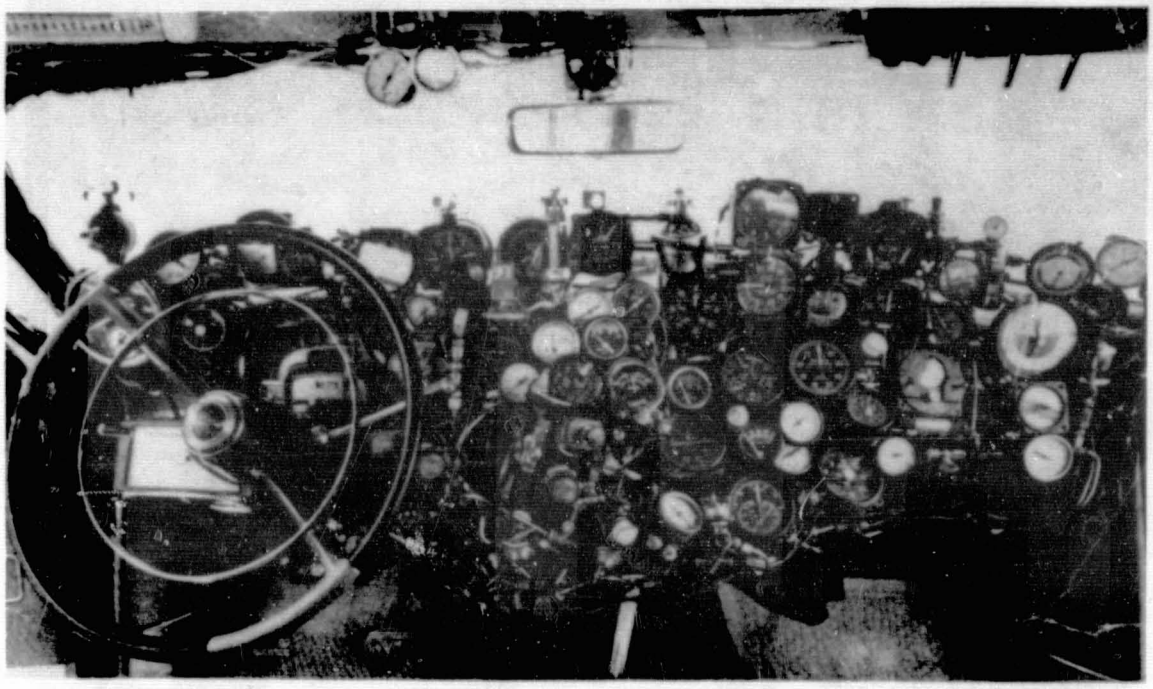


Figure 7. Cockpit of the Future if Present Trend Continues

FLIGHT SIMULATORS

Along with that of the aircraft, the technology of modern flight trainers (simulators) has been evolving for at least six decades. Unlike today's systems which simulate flight and field of vision with realism or fidelity, the early simulation devices were designed primarily to assess the student's suitability for flight training and his resistance to loss of equilibrium (figure 8). As technology advanced, simulators came closer to replicating actual aircraft. A wide variety of sophisticated systems and subsystems became available for integration into training devices. Through incorporation of advanced instructional features such as problem freeze, performance replay, malfunction insertion, and automated performance measurement, today's simulators provide capabilities for improved instruction. In addition to furnishing training and practice in flying maneuvers which cannot or need not be taught in the aircraft, state-of-the-art simulators can be used as teaching tools rather than as substitute aircraft. However, the value of the expected training benefits to be achieved by the inclusion of many of the options in the trainer must be determined.

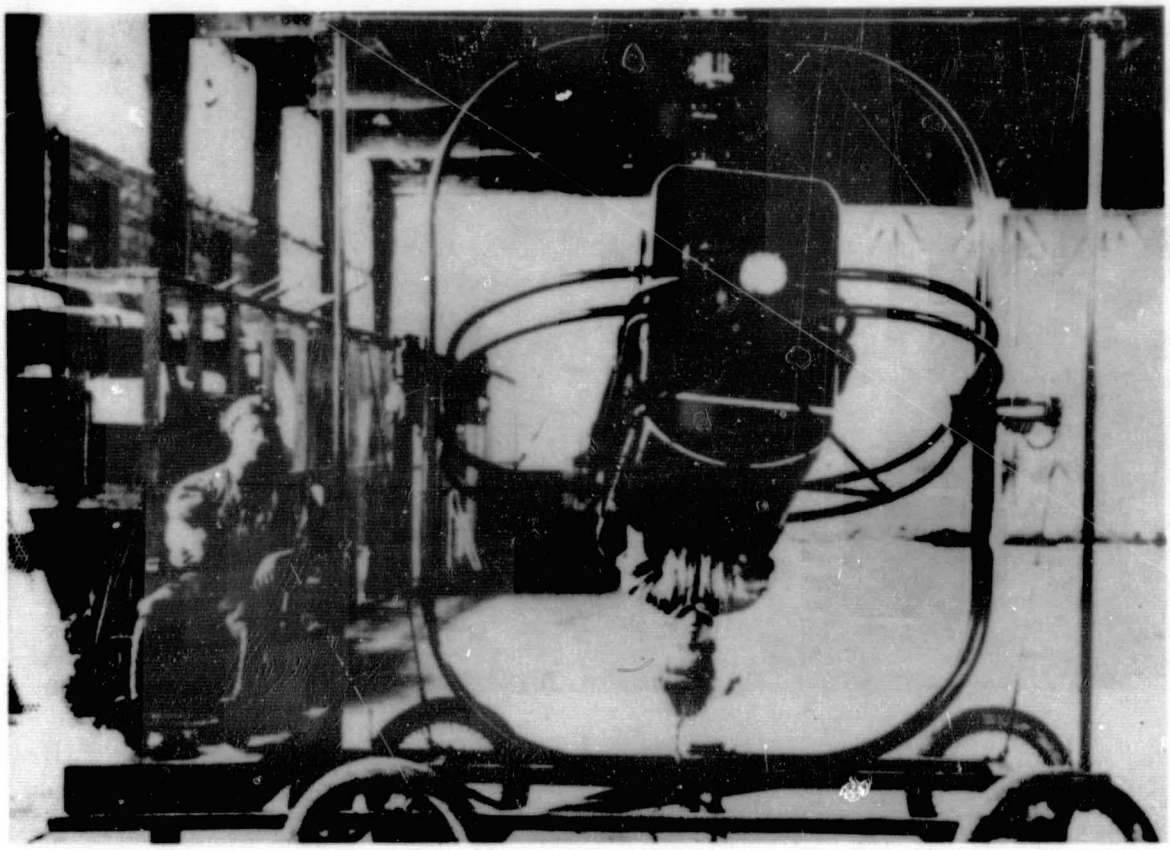


Figure 8. Vintage Trainer

VISUAL DISPLAYS

The development of wide field-of-view visual displays which rely upon control capabilities provided by digital computers has introduced a new area of technology into training simulation. The TV-model board, the oldest of current visual display systems, involves a model terrain board and a TV camera. The camera is driven across the model board by a computer as the aircraft it simulates performs various contact maneuvers. Projected imagery displays involve the use of either a film, a model, or a computer-generated image to furnish the display input which is processed electro-optically and projected onto a curved or dome-shaped screen so that the pilot views the display as he would from the cockpit of an aircraft. Dome-type visuals are used in situations where good quality detail is important and judgements of range/range-rate are important -- air combat maneuvering, formation flight, and aircraft carrier landing training (figure 9).

Computer-generated imagery (CGI) visual systems involve computerized simulation of a visual environment through the use of lines or edges presented on a cathode ray tube or by discrete point light source elements controlled

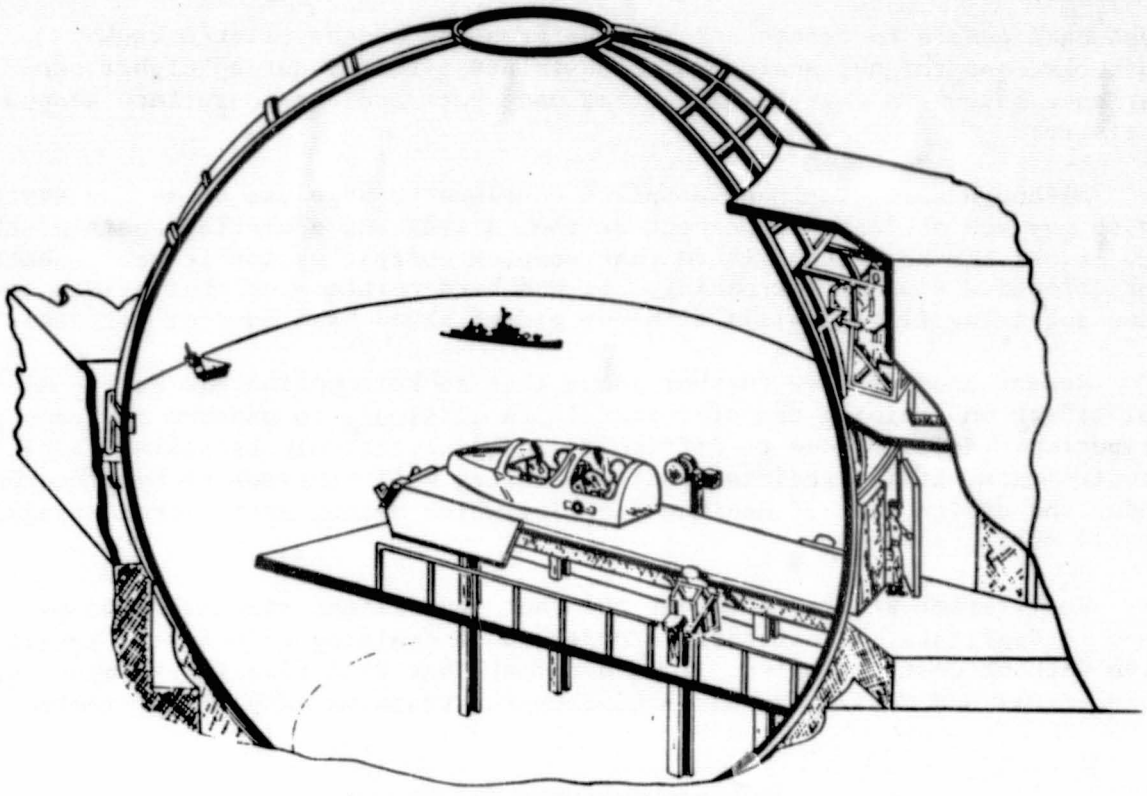


Figure 9. Dome Visual Simulator

to form the in-flight visual environment. Theoretically, this can support instruction and practice of an unlimited variety of training tasks. Dynamic changes in the CGI display are controlled by the simulator's central digital computer.

Cost of current visual simulation systems can exceed six million dollars. Despite this, these systems offer the possibility of substantial cost savings in a pilot's training cycle, especially when replacing those aircraft which have high operating costs. Although there is evidence that positive transfer of training occurs for even the crudest of visual scenes and that visual simulation training aids the student to transition effectively into the airborne environment, studies have also concluded that full fidelity simulation is not necessary for effective transfer of training.

MOTION SIMULATION

Today, there are a variety of devices which can provide motion simulation for a simulator. These include platform motion systems, g-suits and g-seats, stick shaker, and buffet/vibration systems. They are designed to provide either onset or sustained motion cue information to the pilot. The g-seat is an alternative motion cueing system in which air or fluid controls the inflation and deflation of cells in the aircrew seat pan and

seat back panels to create and relieve pressure on the pilot's back, buttocks, and thighs, analogous to sustained g-forces during higher performance maneuvers (air-to-air combat maneuvers and air-to-surface weapon delivery).

Although simulator manufacturers continue to urge use of motion systems which may add at least 10 percent ot the initial and operating costs of the device, it has been established that complex cockpit motion is not essential for effective simulator training. It has been pointed out that pilots have been acquiring flying skills with the aid of fixed base devices for years.

Recent studies have further shown that cockpit motion has such a minimal effect on training transfer that it is difficult to measure its contribution. However, due to difficulty in quantitatively assessing fidelity requirements, it is predicted that simulators will continue to be procured under the design goal of maximum fidelity which means, among other things, costly motion systems.

Most recent studies document the fact that flight simulators do not have to duplicate the aircraft in order to be training effective. However, even without costly studies, it is apparent that full fidelity simulation, both visual and motion, is not necessary for training many flight tasks.

COMPUTER-ASSISTED INSTRUCTION

Today's training buzz word is CAI or computer-assisted instruction. In CAI, all instructional materials, i.e., lessons and tests, are stored in the computer. The student interacts with this material in real time via a terminal and display system. The computer can diagnose student performance, prescribe lessons, maintain records of student progress, and predict individual course completion. A decade ago, it was boldly predicted that CAI was going to revolutionize the learning process. It didn't quite accomplish that. Because it takes an experienced author more than fifty hours of writing to produce one hour of courseware, CAI has yet to develop as much courseware as had been predicted. High hopes have alternated with disillusionment at the unimaginative use of computers for electronic page turning.

We are told that computer cost-effectiveness will double every two years through the 80's. Now the availability of low cost powerful standalone computers has renewed interest in CAI. Other positive elements are the availability of new programming languages, the video disc, greater insight into the learning process, and an understanding of the limitations of computers. There are, however, drawbacks.

The efficacy of CAI for general education has not been demonstrated to be superior to any other method of teaching. In the military, it has been noted that student attrition appears to increase with CAI or with computer

managed instruction. Generally, military instructors are not favorably disposed toward CAI. Studies on the effectiveness of CAI in the military indicate that it saves about 30 percent of the time required by students to complete the same courses taught by conventional instructional methods. However, CAI saves little time over individualized instruction. And, it is estimated that up to 1000 hours of effort must be expended to prepare one hour of courseware for academic material for a flight training program.

In projected flight training programs, more than 30 percent of the student's time in the academic curriculum is spent in front of a cathode ray tube. In addition, from 20 to 25 percent of his simulator training is conducted in a dynamic cockpit procedures trainer which features a giant CRT screen (figure 10). Thus, in this integrated training program, the student pilot spends over 30 percent of his training time in front of a keyboard and one or two display screens. The second display screen is necessary to carry the video disc output which is controlled by the CAI (figure 11). Lessons stored on floppy discs are presented as text and graphics on one display while static and visual depictions are presented by the video disc player on the other display. The CAI provides immediate feedback to the student, and has a built-in performance measurement capability.

TRAINING MANAGEMENT SYSTEM

Each CAI terminal is interfaced by direct data link to a training management system (TMS) which permits two-way information flow. The TMS provides student history, lesson history, identity and availability, scheduling information, student testing, performance measurement, data base interaction, information retrieval, and any number of reports or inputs for reports.

PERFORMANCE MEASUREMENT SYSTEM

In addition to the academic performance capabilities of the CAI and TMS, in-flight performance is measured in the training aircraft by the performance measurement system (PMS). Data gathering is accomplished both by video recording and digital parameter recording. Aircraft flight parameters in specific training situations such as takeoff, approach, and landing, aerobatic maneuvers and instrument flight are recorded. The video recording system records the HUD field-of-view, including HUD symbology and real world view. At debriefing, these cues are presented to the student along with his performance responses which have also been recorded.

Provisions are included in the system for the instructor to call for special snapshot displays to read the value of certain parameters at a specified moment. And, he can also record his evaluation of the student's performance with respect to certain criteria.

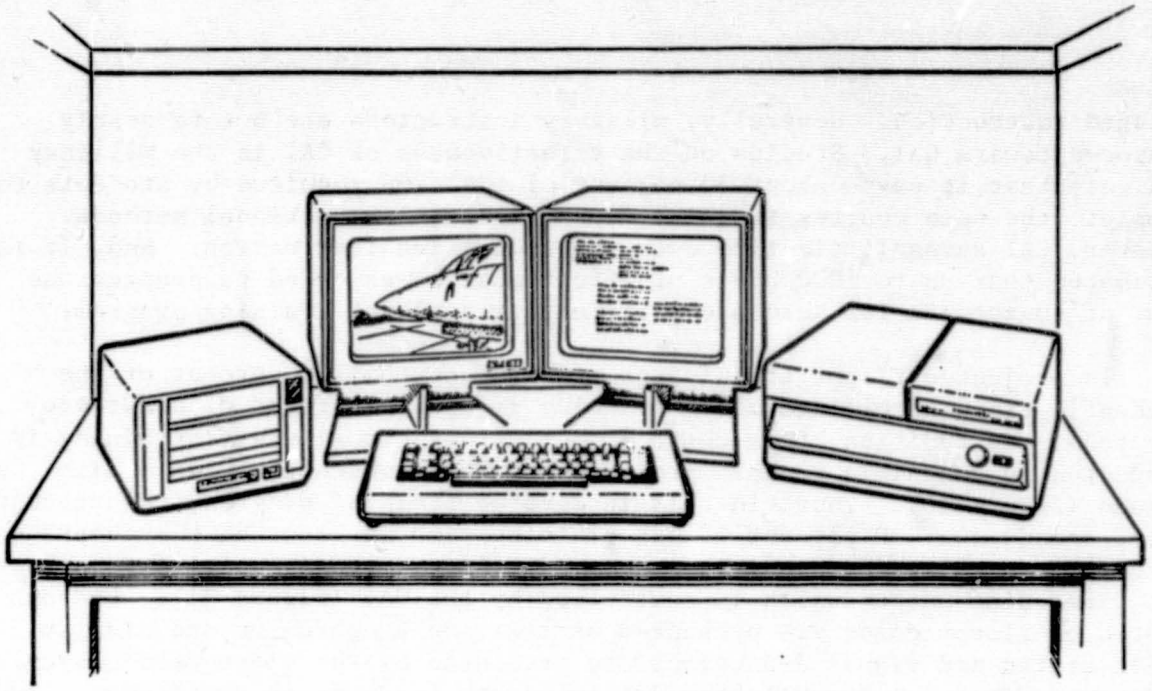


Figure 10. Computer-Assisted Instruction Carrel With Interactive Video Disc

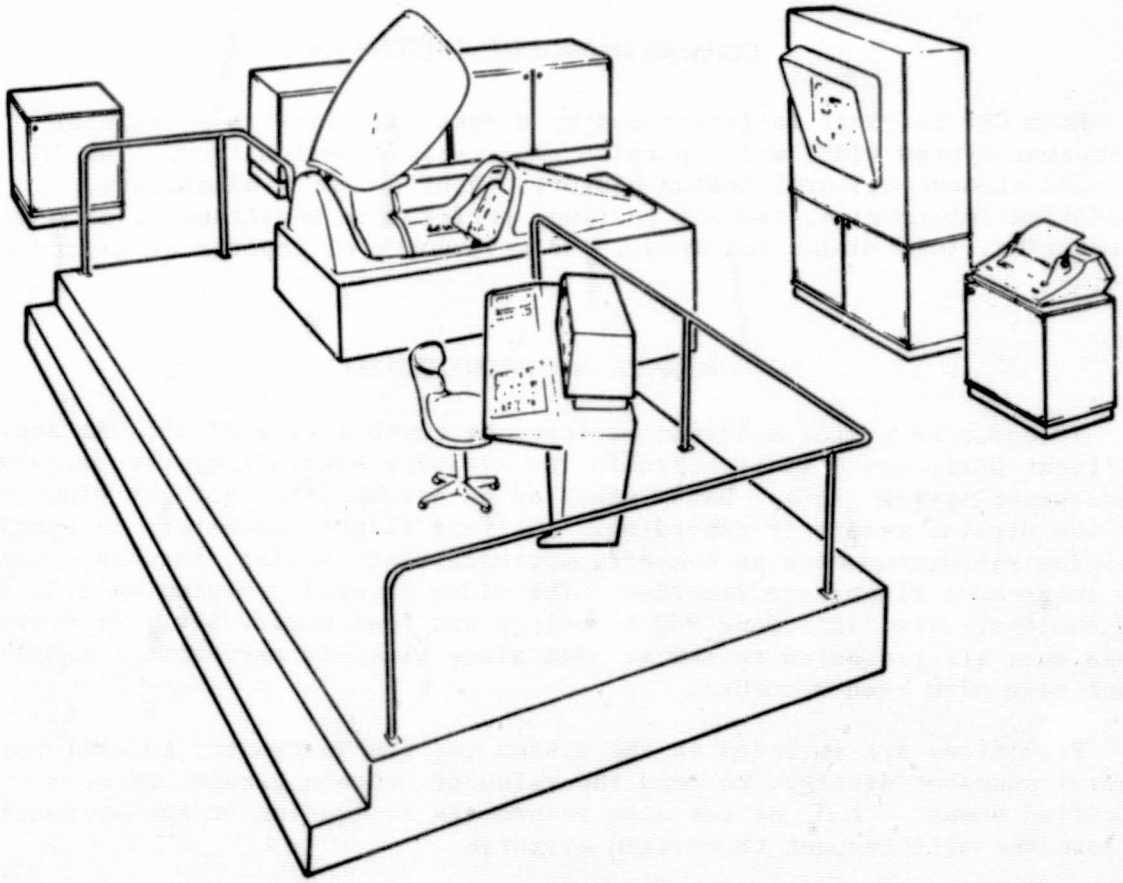


Figure 11. Cockpit Procedures Trainer With Computer-Assisted Instruction

Additionally, all of the performance data recorded for simulator training are interfaced with the TMS. Student performance can be stored in the simulator off-line or real time. The off-line model is of little value for training, however, because the necessity for off-time processing causes feedback delays. On the other hand, real-time measurement offers immediate feedback. However, the requirements for real-time performance measurement implementation are greater because the scoring algorithms must reside on line with the basic simulation program. This requires sufficient memory core and sufficient spare time so that the software can be raised without interfering with the basic simulation program. In addition, sufficient peripherals are necessary so the results can be displayed and stored.

This brief overview, while not addressing, in any depth, the technical intricacies of the TMS and PMS, further exemplifies the multiple man-machine interfaces in pilot training, and the sophistication which may seem to take for granted.

PILOT ACTIVITIES

All human actions can be understood as attempts to achieve a goal; only in this light do they belong together as a related sequence with a definite start and a definite end. To achieve his goal, the pilot engages in two major kinds of activities -- discrimination and manipulation. He must make a series of discriminations -- directional, height or altitude, temporal, and mechanical -- among courses of actions, selecting those which will lead to the accomplishment of a flight mission. The result of these discriminations are subgoals. Manipulation is a psychomotor process which involves moving the airplane's controls in a way that the subgoals needed to execute the mission are accomplished (figure 12).

As part of a system, the pilot does not act in isolation, but receives a continuous stream of inputs in the form of signals, messages, reports, instrument indications, control pressures, and other stimuli. These inputs are combined, interpreted, compared with stored inputs in memory, and transformed into outputs in the form of motor acts and aircraft control movements. More and more, however, the inputs are combined, interpreted, and compared by machine.

The interest of flight training system designers is sometimes linked to statements about how the complexities of modern aircraft cause pilots to make errors. In truth, we are concerned with any pilot error which may occur because of poor training, lack of training, inadequate means for presenting the pilot with information, or controls which are difficult to use. Keeping in mind the cognitively-oriented information processing aspect of the pilot, our interest lies in the fact that the human being is only one element in the system. And, as the trend toward automation continues, the question of man's appropriate role in systems looms large. This has led to consideration of the "Mission Management Control" concept

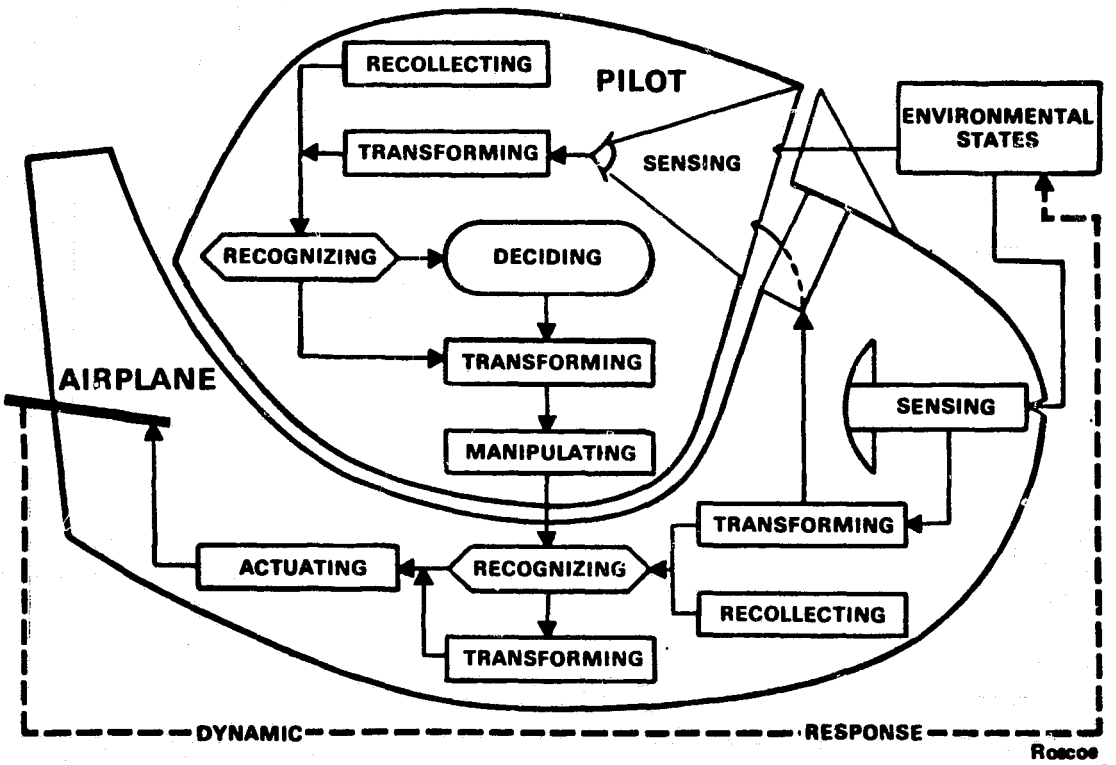


Figure 12. Functional Model of Pilot-Airplane System

wherein the computer presents to the operator a binary "accept" or "reject" solution. The concept essentially takes the operator's brain work out of a problem until the final decision is required. Again, this concept emphasizes the pilot's role as an information manager and final decision-maker.

The pilot in a one-on-one training situation must not be treated as an unthinking, unfeeling, programmed robot. In training, as in any operational situation, we must be certain that man drives the system rather than the other way around. This is the first prerequisite in the process of creating systems and situations for maximum utilization of human talents with corollary human satisfaction in personal accomplishment. Judgement, multimode capability and adaptability, man's contributions to a control system, permit the design of systems with great flexibility and reliability -- and humanity.

CONCLUSION

We have spent considerable time delineating the multiple man-machine interfaces inherent in military jet pilot training. In the search for new and better ways of guiding, monitoring, recording, and evaluating the

student's every move, we have virtually eliminated human instruction. We have shown the complexities of the cockpit and simulator, and examined the attendant stress upon the human information processor. We have addressed the one on one student pilot situation, but much of the same applies to the flight crew situation.

It is obvious that we have come to the end of the streamgage and knob era in aircraft displays and controls. We have entered an era where displays take over more of the routine of mission planning, checklists, enroute checks, etc., and where tactical options and emergency assessments are provided for pilot/operator decision through interactive graphic displays and controls. This shift can be viewed as a change in emphasis from matching man to machine to our developing capability to match the machine to man. Are displays and controls professionals meeting the challenge to exploit this capability at the ultimate interface?

We must make certain that in our hardware hypnosis, in our zeal to provide the best and the most, that we do not denigrate humanity and human interaction, that we do not increase job boredom and reduce motivation and efficienty. Most important, we must not foster within the student a feeling of inadequacy in his ability to cope with problems -- to be in control.

We must capitalize on the computer's capabilities and turn around the thinking of those who find it difficult to cope with inhuman efficiency with which they can neither negotiate nor bargain. Constant and compassionate human interaction can do much to ameliorate the intimidation the student pilot may feel when confronted with a plethora of computers and monitoring equipment.

Above all, we must recognize that more technology is not always better. Considerable unquantifiable human activity requires motivation that is not intrinsic in a machine. And only man himself can insure that motivation.

Psychologists, human factors engineers, and training personnel have great potential to ensure compassionate man-machine interaction. Their design goal should be to take advantage of the behavioral activities of the pilot operators and to circumvent their behavioral limitations in such a manner that motivation and job satisfaction are enhanced. In the end, their greatest challenge is to inject themselves as vital and essential members of the systems engineering team.

REFERENCES

1. Lindahl, Charles E. "Applying Computer Technology to Training Systems." Defense Management Journal, Fourth Quarter, 1980.

2. McCormick, Ernest J. Human Factors in Engineering Psychology and Design (New York: McGraw-Hill Book Co.), 1970.

3. Needham, Richard C., et al. "Flight Simulation in Air-Combat Training." Defense Management Journal, Fourth Quarter, 1980.

4. Perdriel, G., ed. AGARD Conference Proceedings No. 201 on Visual Presentation of Cockpit Information including Special Devices Used for Particular Conditions of Flying (Paris: AGARD), 1976.

5. Roscoe, Stanley N., ed., Aviation Psychology. (Ames, Iowa: Iowa State University Press), 1980.

6. Spooner, A. Michael, et. al. "Visual Simulation in Navy Training." Defense Management Journal. Fourth Quarter 1980.

7. Tomeski, Edward A. and Harold Lazarus. People-oriented Computer Systems (New York: Van Nostrand Reinhold Co.), 1975.

8. Vancott, Harold P. and Robert G. Kinkade. Human Engineering Design (Washington, D. C.: American Institute for Research), 1972.

DECISION AIDS FOR AIRBORNE INTERCEPT OPERATIONS IN ADVANCED AIRCRAFTS

Azad Madni and Amos Freedy
Perceptronics, Incorporated
Woodland Hills, CA

ABSTRACT

Rapid and prompt decision-making during the execution of an F-14 AWG-9 air-to-air intercept mission has been a continuing problem facing the aircrew over the years. The aircrew has had to rely on an inordinate amount of 'gut feel,' rule-of-thumb decisions invariably resulting in ad hoc tactic selection. Consequently, it is generally recognized in the air C^3 community that realtime Tactical Decision Aids (TDAs) are needed by the aircrew in air intercept operations. Fortunately, the extended memory and improved processing capabilities of today's weapon systems computer have made it feasible to incorporate realtime decision aiding algorithms in the onboard software. This paper presents a TDA for the F-14 aircrew, i.e., the NFO (Naval Flight Officer) and pilot, in conducting a multi-target attack during the performance of a Combat Air Patrol (CAP) role. The TDA employs hierarchical multiattribute utility models for characterizing mission objectives in operationally measurable terms; rule-based AI-models for tactical posture selection; and fast-time simulation for maneuver consequence prediction. The TDA makes aspect maneuver recommendations, selects and displays the optimum mission posture, evaluates attackable and potentially attackable subsets, and recommends the 'best' attackable subset along with the required course perturbation.

INTRODUCTION

In a typical F-14 air-to-air mission, the aircrew (Naval Flight Officer and pilot) are called upon to make a multitude of decisions in a rapidly unfolding threat environment. A significant proportion of these decisions impact the overall outcome of the entire mission. Presently, the aircrew make these decisions based upon some combination of training, experience and a limited number of low-level decision aids provided by the F-14's AWG-9 tactical program. These aids, however, have had to be simple due to the limited memory allocation in the onboard computer. However, with the emergence of a large number of sophisticated, high performance threats, it is generally recognized in the air C^3 community that onboard tactical decision aids are required in almost all phases of an air-to-air mission. Fortunately, realtime computer-based TDAs are possible today primarily because of the expanded memory and processing power of today's onboard computers. It is worth noting, however, that despite the evident need, decision aids if not designed from user's viewpoint and task loading can expect great "psychological" resistance from the user pool.

In this paper, a decision aid for assisting the F-14's aircrew in the CAP role of a fleet air defense mission is presented. The aid, based on proven methods from decision analysis, artificial intelligence and fast-time simulation assists the aircrew in the situation assessment and alternative selection functions.

Combat Air Patrol (CAP) Role

In performing its air-to-air missions, the F-14 in the general role of a Maritime Air Superiority Fighter acts as an element of the force Combat Air Patrol (CAP). A typical air-to-air fleet defense mission with the F-14 performing a Combat Air Patrol (CAP) role is given in Figure 1. The CAP objectives are early detection, interception, and attack of airborne threats that endanger the fleet elements. Within the overall CAP role, phases 5, 6, and 7 were selected for aircrew aiding because (1) during these phases the aircrew task loading is high and (2) feasibility of TDAs can be demonstrated in these phases.

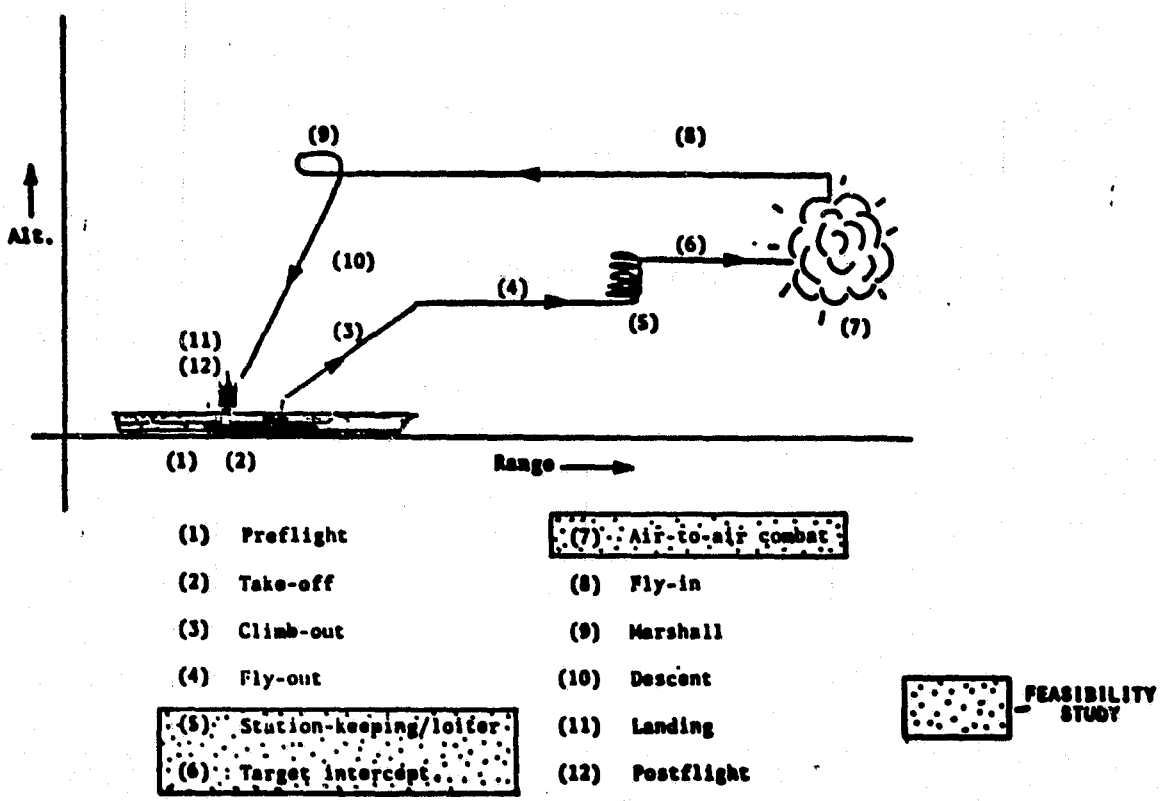


Figure 1. CAP Role Vertical Flight Profile

Phase 5 is station-keeping/loiter. In this phase, the F-14 adheres to a patterned flight at a designated position from the task force. This phase terminates when patterned flight ceases upon target detection.

Phase 6 is target intercept. In this phase, the aircraft pursues a flight path toward a relative position (target conversion) on a selected airborne target. This phase terminates when both the intent to launch a weapon and the capability to effectively launch a weapon exists.

Phase 7 is air-to-air combat. In this phase, the aircraft is flown within a selected weapon launch envelope against a specific target. This phase includes all beyond visual range (BVR) and within visual range (WVR) engagements. The air-to-air combat phase terminates when no further launch capabilities exist or are desired, and the desired return altitude and speed profile has been attained.

Decision Structuring for the F-14 CAP Role

The F-14 CAP role is comprised of a sequence of decisions leading to engagement with and launch on incoming threats. The typical scenario commences with detection and identification of the set of threats. The NFO must quickly switch radar modes, monitor fuel and weapon system status, decide on the subset of the threats to prosecute, select an intercept trajectory, determine the missile launch points, and assess the results.

The sequence of actions in this scenario can be efficiently represented using a decision tree format, as shown in Figure 2. Action decisions, in which the possible choices open to the NFO are listed, are represented by a square box. The decision maker is free to choose only one of the actions. Event nodes, shown as circles, have as branches all the outcomes that may occur at that

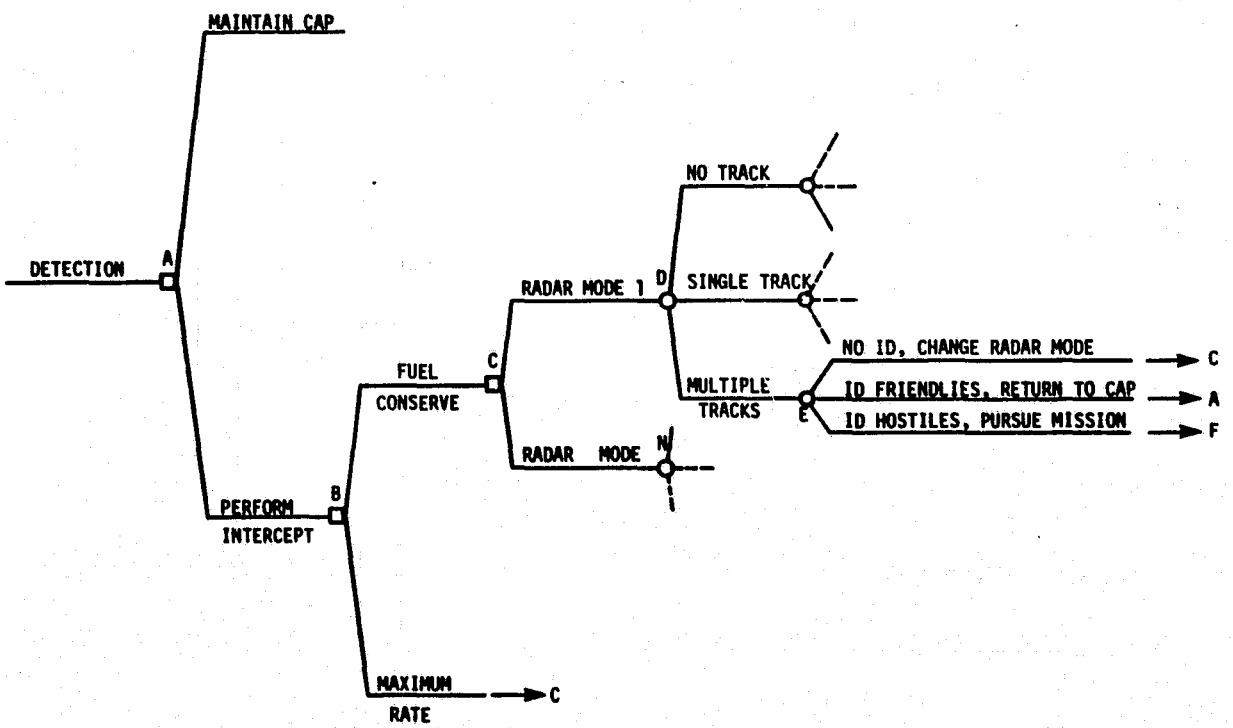


Figure 2. CAP Role Decision Sequence

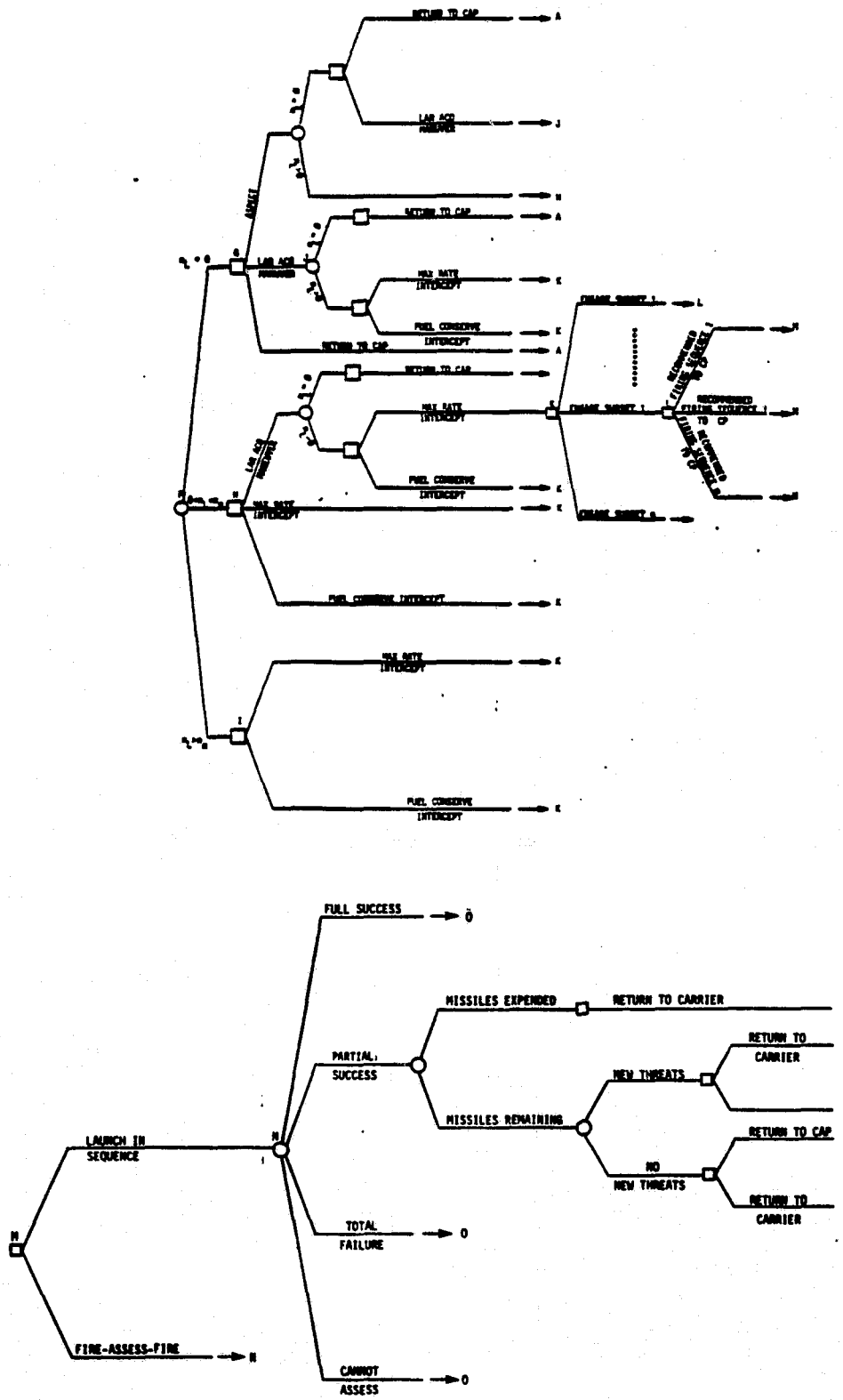


Figure 2. CAP Role Decision Sequence (Cont'd)

point in the tree. The events are characterized by their probability of occurrence and by the value of the outcome. If one path from the beginning of the tree is followed to the end, it describes a possible "scenario." The most effective sequence of actions can be determined by taking the expectation (probability weighting) of utilities over each alternative. The recommended course of action is the one with the highest expected utility.

In order to perform this type of analysis, values and likelihoods must be assigned to each possible outcome. Since there is not enough time to elicit such judgments from the NFO during prosecution of an engagement, results from off-line prior analyses must be loaded into the TDA as routines.

Value estimation in this complex, dynamic environment is best performed using multi-attribute utility (MAU) analysis. MAU methods decompose the complex multi-criterion evaluation problem into more manageable subproblems of scaling, weighting and combining criteria. The MAU evaluation can be expressed as a simple aggregate of constituent factors.

$$\text{Value (Option j)} = \sum_{\substack{\text{events} \\ k}} p(z_k) \sum_{\substack{\text{attributes} \\ i}} \alpha_i U(x_{ijk})$$

where $p(z_k)$ is the probability of occurrence of event k; α_i is the importance weight of attribute i; and $U(x_{ijk})$ is the utility of attribute i associated with option j and event k. This divide and conquer approach of MAU analysis involves defining the problem, identifying relevant dimensions of value, scaling and weighting the dimensions, and finally aggregating the dimensions into a single figure of merit for evaluation. The specific attribute set for evaluation in the F-14 scenario is presented in a later section.

Arriving at estimates of the probability of occurrence of each outcome is also difficult. Two approaches are possible: (1) exhaustively list and estimate off-line the likelihood of each consequence in the CAP scenario or (2) perform fast-time on-line simulations of the maneuver options to analytically determine the major consequences. The first approach, using subjective probability estimates, is expected to be somewhat unreliable and difficult to implement. Even experienced NFOs may be hard-pressed to agree on the probability of adquiring a LAR or encountering a given threat penetration given a specific situation and maneuver. Accordingly, the objectively derived, fast-time simulations were used as much as possible in the TDA operation.

Time Line of the F-14 CAP Role and Aiding Requirements

The current minimally aided F-14 CAP role will be discussed in the following paragraphs with the specific objectives of demonstrating where and when the NFO performance can be enhanced via tactical decision aiding.

The scenario, summarized earlier in Figure 2, commences with the F-14 in a CAP role on the verge of a potential new engagement. The F-14 aircrew are informed of new detections and moments later observe target tracks on the Tactical Information Display (TID). The NFO at this time has to decide if he want to perform intercept or stay on CAP. This decision depends on whether the tracks are identified as 'friendlies' or 'hostiles' and if a successful intercept trajectory to the

oncoming targets is feasible. Currently, target identification once completed, is displayed to him on the TID. However, he has no indication if a successful intercept is possible or not. He draws upon his past experience to make this assessment. If the targets are identified as friendlies, then he stays on CAP. If the targets are identified as hostiles and he decides to embark on an intercept course he has to decide if the intercept should be performed at maximum rate or in fuel conservation fashion. This decision depends on the projected mission profile, and availability of in-flight refueling. Also, he has to determine the details of executing his intercept, i.e., should he "swing" an aspect prior to pursuing an intercept, should he try to acquire Launch Acquisition Regions (LARS) on additional targets or stay with the ones he currently expects to have. Since currently he has no way of knowing what the LAR configuration would be if he executed specific LAR acquisition maneuvers, he makes this determination on his present state of knowledge and 'gut feel.' With regard to performing an aspect he takes into consideration the number of targets he has on the TID, and the number of targets with LARS against his current missile load. Not always will he make the same decision because his perception of secondary factors like time to encounter, and intercept geometry may differ from case to case. However, it is safe to say that if the number of LARS and targets are less than his missile load, he may try to acquire additional LARS. In any case, his next major decision is which subset of targets to go after if there are more targets than missiles and in what sequence to attack them. Currently, the intercept trajectory is usually head collision based on target centroid, emphasizing instantaneous heading and altitude. The firing sequence depends on the order of increasing time until optimum range, nominally fifteen percent into LARS. The current mechanization has manifest drawbacks. There is no objective criteria for attackable target subset selection. The NFO determines who he can go after, generally one at a time, and performs the intercept on that basis. The lead collision intercept trajectory is also suboptimal across the entire spectrum of intercept geometries while the choice of firing sequence is totally ad hoc.

After having "completed" an engagement, i.e., no further attackable threats, the NFO may decide to return to CAP or to the carrier depending on his remaining missile and fuel resources. If he has adequate missile and fuel supply, he prepares for evaluating a reattack if a new wave of threats is detected.

THE TACTICAL DECISION AID (TDA)

Overview

Several key stages of the CAP role immediately present themselves as candidates for aiding. The choice of whether to make an aspect maneuver to gain additional information, what course perturbation to perform to acquire additional LARS, and which subset of threat to engage are all complex decisions well-suited for computer-based aiding. Each of these tasks have well-defined options (turn 15° left, continue on course, etc.) and discrete outcomes (acquire new track, acquire LAR, etc.). Also, the same mission objectives apply to each task.

The portions of the CAP role dealt with by the TDA are summarized in Figures 3 and 4. This tree is roughly equivalent to nodes f through k in the original decision tree (Figure 2). The TDA-assisted decision tree begins with the situation assessment state. After a number of targets are detected and identified, LARS may or not be present on the targets. If no LARS are present, the NFO may elect to stay on the CAP role, prosecute the attack immediately, or perform an aspect

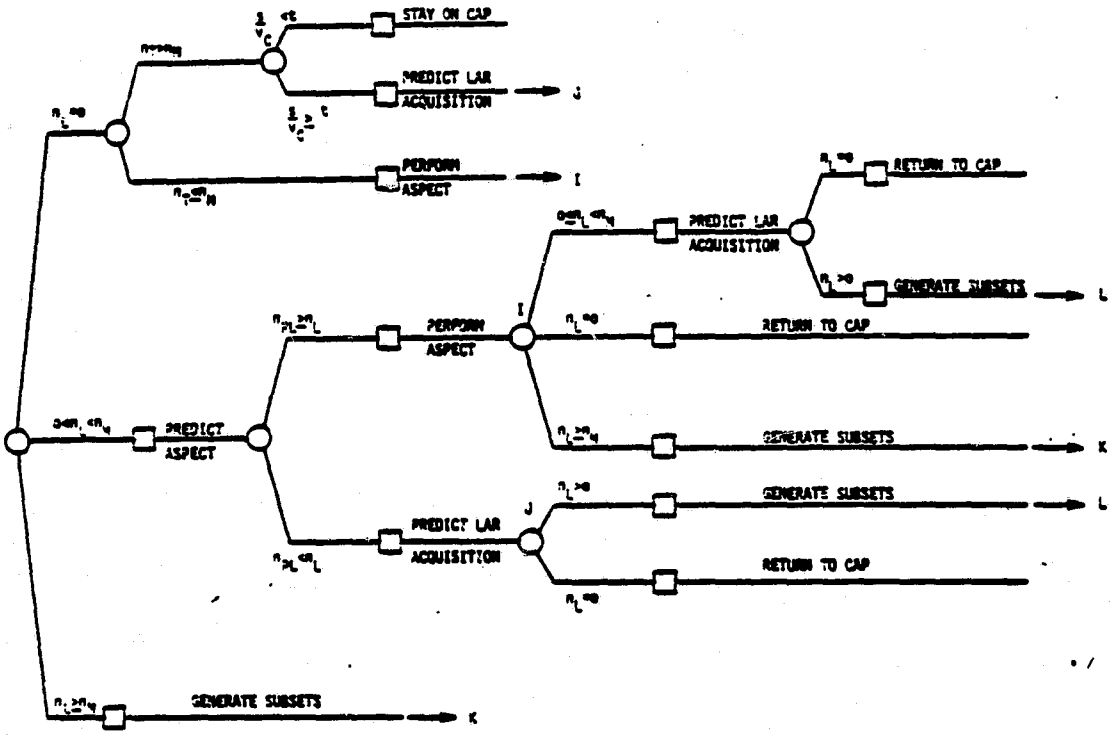
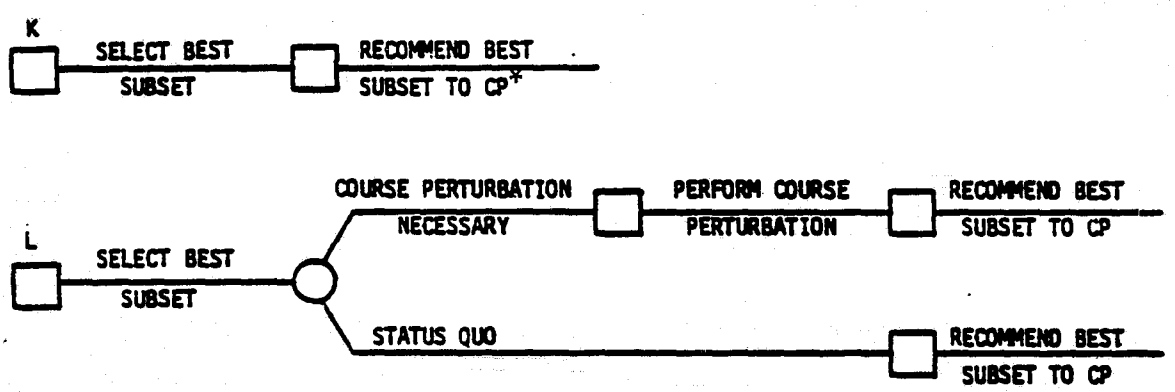


Figure 3. Situation Assessment and Alternative Generation



*CP = Critical point steering algorithm

Figure 4. Alternative Selection

maneuver. The TDA evaluates the options on the basis of the number of targets present (n_T), the number of missiles onboard the F-14 (n_M), and the time to engagement (t_E): and make a recommendation to the NFO. Similar situation assessments are made if there are targets with LARS at the initiation of the engagement. Of course, the criteria of evaluation employed in the actual TDA evaluation are more complex than that described above. The TDA considers the impact of each choice and outcome on the threat to the carrier, on the damage inflicted on the enemy, and on the F-14's own vulnerability. The specific criteria are developed in detail in subsequent sections. The next stage in the decision tree leads to alternative generation (nodes I and J in Figure 3).

If an aspect maneuver is recommended, on the basis that the predicted number of LARS following the maneuver (n_{PL}) is greater than the original number of LARS (n_L), then following the maneuver, the NFO must choose to continue prosecution of the threats or return to CAP. A return to CAP would only be called for if following the maneuver no LARS were present.

The above sequence illustrates an important characteristic of the TDA. Instead of requiring subjective estimates of the likelihood of LAR acquisition in each situation, the LAR tests are made by calling a fast-time simulation. In this way, the decision aiding is based on hard data of position, course and speed of the F-14 and the threats.

Once the aspect maneuver is complete, course perturbation checks and subset generation are performed by the TDA (Nodes K and L). Here changes in heading, altitude and speed are tested to see if additional LARS result. Then the "best" threat subset is recommended for attack. The specifics of what is "best" will be covered in the next section.

In the following paragraphs, the structure and operation of the TDA will be presented in terms of: the mission objectives hierarchy which "drives" the aid, the automated programs for mission posture specification, aspect maneuver recommendation, course perturbation, target subset selection, and display requirements.

Objective Structuring and Mission Success Hierarchy

The overall mission objective for the F-14 CAP role starting with target detection and culminating with target reattack can be summarized in three key tradeoff objectives: (1) maximize carrier safety; (2) maximize tactical gains; (3) minimize resource expenditure. Each of these objectives can be embedded in a linear multi-attribute representation framework and can be further decomposed into explicit sub-objectives that themselves constitute measurable attributes or have measurable attributes associated with them. Each of these attributes provide a scale for measuring the degree of attainment of the associated sub-objective. The weighted combination of these attribute levels provide an indication of the attainment of each parent key objective. The weighted combination of the level of attainment of each key objective then provides a measure of the overall mission success objective. The mission success hierarchy is shown in Figure 5. The actual choice of the attribute set is extremely important. Dawes (1974) states that the choice of factors to include is probably of greater impact than the determination of the model form. Desirable characteristics are

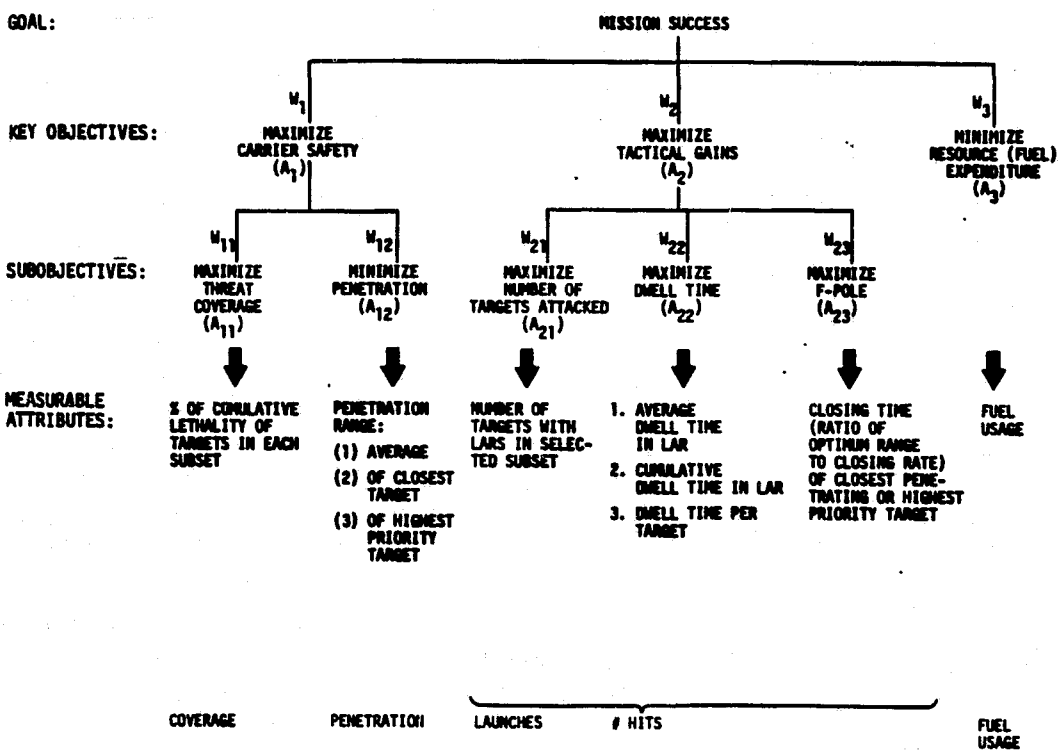


Figure 5. Mission Success Hierarchy

accessibility for measurement, independence, monotonicity with preference, completeness of the set, and meaningfulness for feedback. Monotonicity, in this content, implies that an increase in the attribute level always results in an increase in preference. If the attribute levels are monotonic, a simplification is possible. Fisher (1972) and Gardiner (1974) note that a straight line approximation to the utility function results in minor losses of model accuracy. The attributes selected within the framework of the three key tradeoff objectives that characterize the F-14 CAP role mission phase possess the desired characteristics described above. These attributes were elicited from Naval Flight officers and pilots who jointly agreed upon the selected attribute set.

The first key objective, maximizing carrier safety, can be decomposed into maximizing threat coverage and minimizing target penetration. Threat coverage is measured in terms of the threat associated with the engaged subset of targets. The greater the threat engaged the higher the threat coverage. Target penetration range is defined as (1) the range from Task Force Center (TFC) of either the closest penetrating target attacked or the highest priority target attacked (whichever is chosen due to situation). Minimizing target penetration range is equivalent to maximizing the range from the TFC of either the closest penetrating or highest priority target while ensuring that the average range of the remaining targets from the TFC is above a predeterimined range threshold.

Maximizing tactical gains is analogous to maximizing expected kill (E_K). This objective is not amenable to direct measurement but can be decomposed into related operationally measurable objectives, the attainment of which implies the attainment of the related parent objective. Thus, E_K is expressed in terms of maximizing (1) the number of Phoenix launch opportunities or the number of targets attacked, (2) the dwell time in LAR, and (3) F-pole. The number of Phoenix launch opportunities is the number of LARS the F-14 obtains on the threat subset plus the number of second shot opportunties. A second shot opportunity is predicted if a target previously acquired is expected to have a LAR at a time $t > k$ later. The second attribute, number of targets attacked, is the number of distinct targets in the subset upon which LARS are predicted. The dwell time in LAR is the predicted time in seconds between the entry and exit points summed across all LARS in the subset. The final attribute, F-pole is not directly predictable. However, it is proportional to R_{OPT} and inversely proportional to the closing rate, V_C. Thus, R_{OPT}/V_C is employed as an indicator of F-Pole.

The final key objective, minimizing resource expenditure, implies minimizing fuel expenditure. The predictable attribute corresponding to fuel expenditure is the fuel remaining after each of the perturbation maneuvers. The prediction of each of the operationally measurable attributes is discussed below.

Predicted Attribute Level Computation. Predictive computation of the various attributes that define mission success hierarchy are defined on normalized 0 to 1 scales.

1. Threat Coverage. Threat coverage is defined as the weighted sum of the total number of attackable targets (with LARS) in the subset of targets selected for engagement. The weighting factor associated with target i is its lethality index, i.e., the lethality of target i. Lethality, in general, is primarily a function of the onboard weapon load and the EW capability of that target. These two parameters can usually be determined once the target has been identified. There are also some generic tactical doctrines that drive the lethality computation. For instance, it is generally agreed upon by the operational community that platforms should be attacked first, so that they cannot return another day and pose a recurring threat to the NFO. Additionally, attacking platforms first provides a tactical advantage in that the platforms are denied midcourse guidance correction. A second doctrine is that manned aircrafts should be attacked/engaged prior to attacking any missiles. However, since such detailed lethality indices were not available during the study, a priori lethality values were assigned to the different targets modeled in the multi-target KIWI simulation environment. The initial implementation was in the form of a table look-up of lethality index versus target type for the candidate threats that were simulated. With this simplification, threat coverage can be computed as:

$$\text{Threat Coverage} = \sum_{i=1}^{k} \ell_i$$

where ℓ_i, i=1, ... k, is the lethality of target i; k is the total number of targets in the subset selected for attack.

The above computation is normalized relative to the product of the number of missiles currently onboard and the lethality index associated with the most lethal target, Thus,

$$a_{11} = \sum_{i=1}^{k} \ell_i / N_{miss} \cdot \max_i (\ell_i)$$

2. Penetration Range. Penetration range is defined as a weighted combination of the distance from the Task Force Center (TFC) of either the closest penetrating target attack or the highest priority target attacked, and the average range from TFC of the remaining targets on the TID. Penetration range as defined here should be maximized for the successful attainment of mission objectives. If $\underline{r}_T$ is the location (position vector) of the closest penetrating target or highest priority target (depending on the context), $\underline{r}_{TFC}$ is the position vector of the Task Force Center and $\underline{r}_C$ is the position vector of the centroid of the remaining targets, then maximizing penetration range implies maximizing

$$a'_{12} = (1-\epsilon)\underbrace{|\underline{r}_T - \underline{r}_{TFC}|}_{\text{primary objective}} + \epsilon\underbrace{|\underline{r}_C - \underline{r}_{TFC}|}_{\text{secondary objective}}$$

1-ε and ε are the weights associated with the primary and secondary objectives, respectively.

$\underline{r}_C$ is defined by

$$\underline{r}_C = \frac{1}{n-1} \sum_{i=1}^{n-1} \underline{r}_i$$

where n is the total number of targets on the TID and $\underline{r}_i$ is the location of target i. Normalizing, attribute

$$a_{12} = a'_{12} / |\underline{r}_{max} - \underline{r}_{TFC}|$$

where $\underline{r}_{max}$ (= $\max_i \underline{r}_i$) is the position of vector of the farthest target in the threat cloud.

3. Number of Phoenix Launch Opportunities or Number of Targets Attacked. The number of Phoenix launch opportunities can be predicted conservatively on the basis of the number of targets (i.e., LARS) in the selected subset. This definition, of course, assumes that no existing LARS will be lost nor new ones acquired during the course of the impending engagement. Thus, the number of Phoenix launches, n, is given by

$$\boxed{\begin{aligned} n &= n_T, \text{ if } n_T \leq n_M \\ n &= n_M, \text{ if } n_M > n_T \end{aligned}}$$

where n_T is the number of targets in the subset and n_M is the onboard missile load.

The normalized attribute is then given by

$$\boxed{a_{21} = \frac{n}{n_M}}$$

4. Cumulative Dwell Time. Dwell time, t_D, is defined as the time spent in the LAR of a given target. Thus, dwell time in LAR of target i, t_D, is given by

$$\boxed{t_D^i = \frac{|R_{max}^i - R_{min}^i|}{V_C^i}, \quad i=1, 2, 3, \ldots, k}$$

Cumulative dwell time, T_D, is the sum of the dwell times in the launch zone of each individual target.

$$\boxed{T_D = \sum_{\text{target } i} \frac{|R_{max}^i - R_{min}^i|}{V_C^i}, \quad i=1, 2, \ldots, k}$$

The normalized attribute dwell time (a_{22}) is given by

$$\boxed{a_{22} = T_D \Big/ K\alpha}$$

where $\alpha = \max_i \left[t_D\right]$

5. F-Pole Range. F-Pole range is the fighter to target range at the end of the predicted missile TOF. Maximizing F-Pole range is equivalent to maximizing closing time for that target. The closest penetrating target i in each subset k is found from the closing time to the TFC, $t^k_{c_i}$.

$$t^k_{c_i} = \frac{|\underline{r}^k_i - \underline{r}_{TFC}|}{V^k_{c_i}} = \min_j \; t^k_{c_j}$$

where $V^k_{c_i}$ = closing rate of target i in subset k on TFC = $\underline{U}^T_{LOS}|\underline{V}_i - \underline{V}_{TFC}| \approx \underline{U}^T_{LOS}\,\underline{V}_i$

$\underline{r}_{TFC}$ = position of vector of TFC

$\underline{r}^k_i$ = position vector of target i in subset k

$\underline{V}_i$, V_{TFC} = velocities of target and TFC, respectively

For each closest penetrating target i in each of the subsets k, compute

$$\alpha_k = \left(R_{opt_i} / V_{c_i} \right)_k \quad ; \quad \alpha_{max} = \max_k \; \alpha_k \quad ; \quad k = 1,2, \ldots .;$$

$R_{opt_i} \approx 85\ R_{max_i}$ = optimum range; V_c = closing rate between fighter and target = $\underline{U}^T_{LOS}|\underline{V}_i - \underline{V}_F|$

The normalized attribute level, a_{23} for each subset k is given by

$$a_{23} = \alpha_k / \alpha_{max}$$

6. Fuel Usage. The last attribute that has to be predicted is the fuel remaining following a perturbation maneuver. Each perturbation maneuver requiring a change in altitude, speed or heading can be ranked in terms of fuel requirements from the most fuel intensive to the least. The ranking along with the fuel requirements on a scale of 0 to 1 is given in Table 1. The attribute level derives directly from the table.

Table 1. Fuel Requirements of Perturbation Maneuvers

PERTURBATION MANEUVER	FUEL REQUIREMENT
1. CLIMB (+ 2000 Feet)	1.0
2. INCREASE SPEED (+ .1 MACH)	.8
3. CHANGE HEADING (+ 10°)	.2
4. DESCEND (- 2000 FEET)	0
5. DECREASE SPEED (- .1 MACH)	0

Situation Assessment Aid

Situation assessment in a tactical environment is the process of determining the values or levels of the salient attributes or dimensions that characterize the tactical problem confronting a decision maker. In the F-14 CAP role aiding context, the situation assessment aid provides the NFO with prompt and timely estimates of the tactical situation confronting him. This "sensed" information enables the NFO to maintain intimate contact with the time-varying data that characterizes the relevant dimensions of the tactical environment. This aid recommends a suitable mission posture to the NFO based on a combination of internal and external "sensed" conditions. It, further, evaluates if an aspect maneuver is warranted given the prevailing tactical configuration and recommends an aspect if it is indicated. In the following paragraphs, the key mission postures will be identified along with a set of conditions that exhaustively span the transitions from one posture to the other. Included also is the rationale and criteria for performing an aspect maneuver.

Tactical Mission Postures. The relative weighting on the various attributes in the mission success hierarchy varies according to the tactical situation. A total of six postures have been identified, each with a distinct set of attribute weights:

(1) Offensive -- maximize number of enemy downed, with secondary goals of maximizing carrier safety and resource conservation. ($W_2 >> W_1, W_3$) where W_1, W_2, W_3 are shown in Figure 5.
(2) Defensive -- maximize carrier safety, with secondary goals of maximizing E_k and resource conservation. ($W_1 >> W_2, W_3$)
(3) Conservative/Offensive -- maximize E_k and resource conservation. Virtually ignore carrier safety. ($W_2, W_3 >> W_1$)
(4) Conservative/Defensive -- maximize carrier safety and resource conservatioon. Virtually ignore E_k. ($W_1, W_3 >> W_2$)
(5) Carrier Safety -- maximize carrier safety alone. ($W_1 = 1; W_2 = W_3 = 0$)
(6) E_k -- maximize E_k alone. ($W_2 = 1; W_1 = W_3 = 0$)

The first four postures, Offensive (O), Defensive (D), Conservative/Offensive (C/O), and Conservative/Defensive (C/D), are "trade-off" strategies. Different combinations of attributes are emphasized in each. In the offensive posture, for

instance, E_k is emphasized at the expense of carrier safety and resources. The final two postures, Carrier Safety and E_k are "pure" strategies. Carrier safety puts zero weighting on E_k and resources. E_k only weights the E_k attributes.

The six postures correspond to distinct tactical situations. These can be classified by conditions associated with the following tactical variables:

(1) Threat penetration range. The threat distance from (a) the task force center or (b) the weapon release line around the carrier.
(2) Fuel remaining. The amount of fuel left to return to the carrier.
(3) Numerical advantage. The numbers of missiles compared to the number of targets.
(4) Lethality. Phoenix missile-equivalents of the threats.

Posture Transition Criteria. An exhaustive set of relations between the postures and the conditions are given in Table 2. For example, defensive posture (P_2) is called for if high threat level is present, sufficient fuel remains, either a numerical advantage or disadvantage exists, and low lethality is present. Similar descriptions for the choice of the other postures can be given (see Table 2).

Table 2. Posture Transition Logic

	Meaning	C_1 C_2 C_3 C_4	Transition Logic
P_1	Offensive	0 0 1 1 0 0 0 -	If $(\overline{P}_1 \wedge (\overline{C}_1 \wedge \overline{C}_2) \wedge ((C_3 \wedge C_4) \vee \overline{C}_3))$ Then P_1
P_2	Defensive	1 0 - 0	If $(\overline{P}_2 \wedge (C_1 \wedge \overline{C}_2 \wedge \overline{C}_4))$ Then P_2
P_3	Conservative/ Offensive	0 1 1 -	If $(\overline{P}_3 \wedge (\overline{C}_1 \wedge C_2 \wedge C_3))$ Then P_3
P_4	Conservative/ Defensive	0 1 0 - 1 1 - -	If $(\overline{P}_4 \wedge C_2 \wedge ((\overline{C}_1 \wedge \overline{C}_3) \vee C_1))$ Then P_4
P_5	Carrier Def. Only	1 0 - 1	If $(\overline{P}_5 \wedge (C_1 \wedge \overline{C}_2 \wedge C_4))$ Then P_5
P_6	EK	0 0 1 0	If $(\overline{P}_6 \wedge (\overline{C}_1 \wedge \overline{C}_2 \wedge C_3 \wedge \overline{C}_4))$ Then P_6

CONDITIONS:

C_1: Enemy poses high threat, i.e., at least one target close to WRL; $\overline{C}_1$ = low threat

C_2: Fuel critical, i.e., insufficient fuel to complete engagement and return to carrier; $\overline{C}_2$: fuel critical

C_3: Numerical advantage, i.e., onboard missile load $\geq$ number of targets; $\overline{C}_3$ = numerical disadvantage

C_4: Cumulative lethality, L, of subset high, i.e., $L \geq L_{min}$; $\overline{C}_4$ = low lethality

Transition-Conditions. There are four conditions which are monitored to determine posture transitions. These conditions are: C_1 – threat level, C_2 – fuel status, C_3 – numerical advantage status, and C_4 – cumulative lethality.

There are two values necessary to compute the C_i's. These are LMIN and FUEL-THRESH. LMIN is the threat cloud lethality threshold (based on a critical missile load).[1] This value is contingent on the scaling of the lethality values that reside in a table look-up. Since lethality values associated with different targets were unavailable, each target in KIWI was assigned a lethality value. This information was contained in a table look-up. More sophisticated lethality computation scheme based on target identification and missile carrying capacity can replace this table look-up in a straightforward manner. FUEL-THRESH is the amount of fuel required on the average to fire all missiles onboard the F-14 and return to the carrier. THRESH2 is the minimum time threshold for the closing time between threat cloud centroid and weapon release time. With FUEL-THRESH calculated for the current situation, the C_i's can be computed as follows:

(1) If the closing time for the centroid of the threat cloud to weapon release time is less than some threshold THRESH2, set C_1 to high; otherwise, set C_1 to low.
(2) If the fuel remaining is less than FUEL-THRESH, the critical fuel threshold, set C_2 to high; otherwise, set C_2 to low.
(3) If the onboard missile load is less than the number of targets, C_3 is set to low; otherwise C_3 is set to high.
(4) If the cumulative lethality of the current threat cloud is less than LMIN, C_4 is set to low; otherwise, C_4 is set to high.

The mapping of the tactical conditions to postures allows the automated transition from one posture to another as the sensed situation changes. The TDA has access to all of the condition levels and contains the logic for transitioning between postures.

Aspect Maneuver Recommendation. The first point at which the TDA aids the NFO is in the initial aspect maneuver. Here a set of threats has been detected, and a decision is needed on whether to perform a horizontal aspect maneuver to resolve additional threats behind those currently being tracked. The disadvantage of an aspect maneuver is that the F-14 may end up with an inferior tactical position, i.e., have less LARS than currently predicted. What the NFO needs is some means of predicting what the resultant LAR configuration would be if he performed an aspect maneuver. The TDA performs this predictive computation. If the number of LARS is strictly less than the number of missiles but greater than zero, the TDA predicts the consequences of a canned aspect maneuver. If the number of LARS expected after performing aspect is greater than or equal to the number currently expected, the TDA displays the aspect recommendation on the TID, by changing the steering dot and displaying "ASPECT" just below the TID buffer readout. If the number of LARS is zero and number of targets is less than or equal to the number of missiles then aspect is predicted and recommendation displayed. On the other hand, if $n_t > N_m$, the F-14 either stays on CAP or tries to acquire LARS depending on the duration of time to encounter.

[1] I.e., LMIN = $\sum$ lethality i, for all i, where the summation is for each target in the threat cloud.

The NFO may override the aspect recommendation if he acquires external knowledge of the raid structure or if he spots a high priority target in the process. If such information is available, he overrides the aspect recommendation by pressing the "OVERRIDE" button on the TDA-dedicated portion of the CAP panel. This action results in returning the steering dot to the original position and blanking the readout. If the NFO decides to go along with the recommendation, the aspect maneuver is performed. (The actual number of resultant LARS may or may not be the same as the number predicted.) The aspect maneuver itself is set to be a 20-30 degree heading change in the horizontal plane in the direction of increasing aspect. The time to complete the maneuver is established in a look-up table.

The new coordinates are calculated by the TDA and fed to the steering dot program. If the aspect maneuver is recommended, aspect is shown on the TID drum and the steering dot is moved to the new course.

During the time required for the aspect maneuver, the TDA programs are suppressed. This is because a sequence of targets with or without LARS may come into view during the maneuver, requiring continuous updating. Instead, a prediction of time to aspect completion is computed and all processing halted until that time has elapsed.

If performing the aspect maneuver results in total loss of LARS and time to encounter is less than time to perform LAR acquisition, the NFO may wish to return to the CAP role; otherwise, he tries to acquire LARS. In any case, the next major decision is one of acquiring additional LARS and generating attackable subsets versus generating attackable subsets directly.

A special situation is present if more than nine targets appear initially. Because of computational constraints, large target sets are pruned down to the nine most important targets and these targets are then considered by the TDA. The measure of selection is the lethality (missile carrying capacity) divided by the closing time (distance/closing velocity). The threats are ranked in decreasing order of this measure and the top nine selected for TDA processing.

Aspect Prediction. Aspect Prediction calculation is called for prior to making an aspect recommendation. Typically, an aspectmaneuver is made in order to resolve the individual targets in a multiple target raid. As aspect maneuver is culminated when the aspect angle is about 30^{o}. Where exactly the fighter ends up at the end of an aspect is a function of the F-14 and target heading angles, the current F-14 and target velocities, the nominal g-forces employed in a turn and the basic assumptions associated with the turn. For instance, a constant velocity, fixed radius turn assures minimum fuel usage while a move to the desired aspect in minimum time assures the performance of the maneuver in an acceptable amount of time. In this instance, each of the considerations are warranted because fuel should be conserved when possible and the meneuver should be done as fast as possible to allow the NFO adequate time to evaluate TDA recommendations en route to actual engagement. The aspect prediction equations are derived in Madni, et al (1980).

Alternative Generation and Selection Aid

Attackable Target Subsets Generation. An attackable subset is defined as a group of targets in which all targets have LARS for a specific location of the F-14. Incremental changes in the F-14 location can result in a totally different attackable subset. An attackable subset can be no smaller than the onboard missile load if the number of targets is greater than the on-board missile load. Thus, if the onboard missile load is four, say, then the attackable subset can have no less than four targets.

If the number of LARS is greater than or equal to the number of missiles, the attackable subsets are generated directly by forming all feasible combinations of targets with LARS corresponding to the current state vector. If the number of LARS is less than the remaining number of missiles onboard and time to encounter is greater than or equal to time to LAR acquisition, then LAR acquisition maneuvers are predicted. All possible attackable subsets are generated by forming all feasible combinations of targets with LARS corresponding to the current and perturbed F-14 state vectors. If in this process, a priority target is selected by the NFO by pressing PRIORITY function button on the CAP panel followed by manually hooking the target on the TID, then all possible attackable target subsets are scanned to determine if they contain the priority target. Only those subsets that contain the priority target are viable candidates for subsequent evaluation.

Attackable Subset Selection. The subset evaluation is accomplished by first forming all feasible combinations of targets with LARS. Feasibility demands that the number of targets in the subset is less than or equal to the number of missiles onboard and that all geometric constraints are satisfied. Then each subset is assigned a vector of attribute levels according to projected performance. The attributes are defined as before in Figure 5. In this mission success hierarchy all measurements are normalized to 0 to 1 scales, where zero corresponds to the worst case possible (max lethality coverage). The attributes are weighted by importance. The weighting is also normalized so that an overall zero implies zero on all attributes and an overall one implies a one on all attributes. The relative weighting itself differs according to tactical situation (the postures in Table 2) and is estimated through expert elicitation.

Course perturbation takes the form of positive and negative changes in heading, speed and altitude. The step size is defined at $\pm$10 degrees in heading, $\pm$.1 Mach, and $\pm$2000 feet in altitude. All perturbations are checked to determine the number of LARS present. Target subsets for each perturbation are evaluated according to the above process (Figure 6). In the end, a maneuver and subset is recommended which has the highest overall utility, according to the mission posture and conditions. This recommendation is displayed by showing "subset recommendation" just below the TID buffer readout, and moving the steering dot on the TID display.

If the "Priority" button is pressed, the priority target can be manually hooked by the NFO on the TID. All subsets containing the priority target are evaluated as before with the subset with the highest overall utility under the prevaility posture being recommended. The manual subset select button allows the NFO to disengage the TDA and hook the desired target set manually. The selected targets are output directly to the steering algorithm.

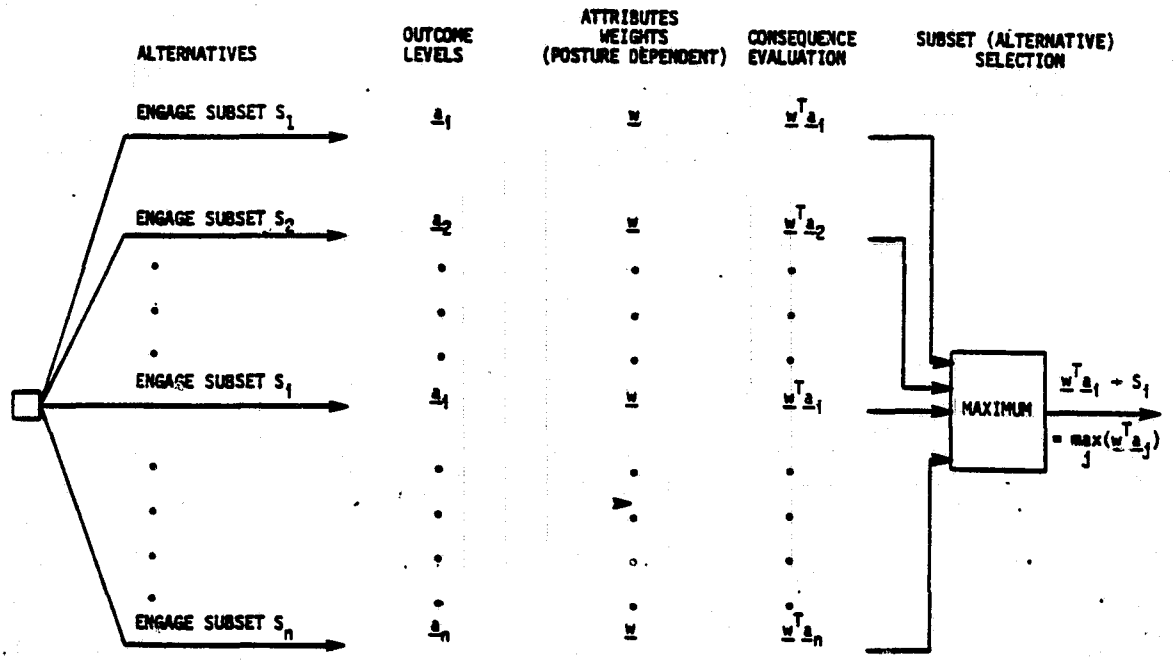


Figure 6. Subset Selection

NFO-TDA Interface

CAP Panel Modification. At the very outset of this project, it was established that the Computer Access Panel (CAP) was going to be used by the NFO to communicate his inputs to the TDA software and that TDA-related information was going to be displayed to the NFO on the Tactical Information Display (TID). Consequently, the NFO-TDA interface was configured to fit within the space and configurational constraints of the aforementioned devices.

The TDA's implementation on the CAP panel was determined after extensive discussions with Naval Flight Officers (NFOs). The implementation impacts the CATEGORY switch and the message selection pushbuttons. The CATEGORY switch is a six position switch which permits sharing the message selection pushbuttons. When the CATEGORY switch is rotated, a matrix of labels next to the MESSAGE pushbuttons is changed. When the CATEGORY switch is set to TDA, the legend associated with the ten multi-purpose MESSAGE pushbuttons correspond to the posture, target and override controls (Figure 7). This mechanization was selected on two counts: (1) in the opinion of the NFO's, it fit naturally within their task structure and user interface, and (2) it resulted in minimal hardware changes to the existing controls and display configuration.

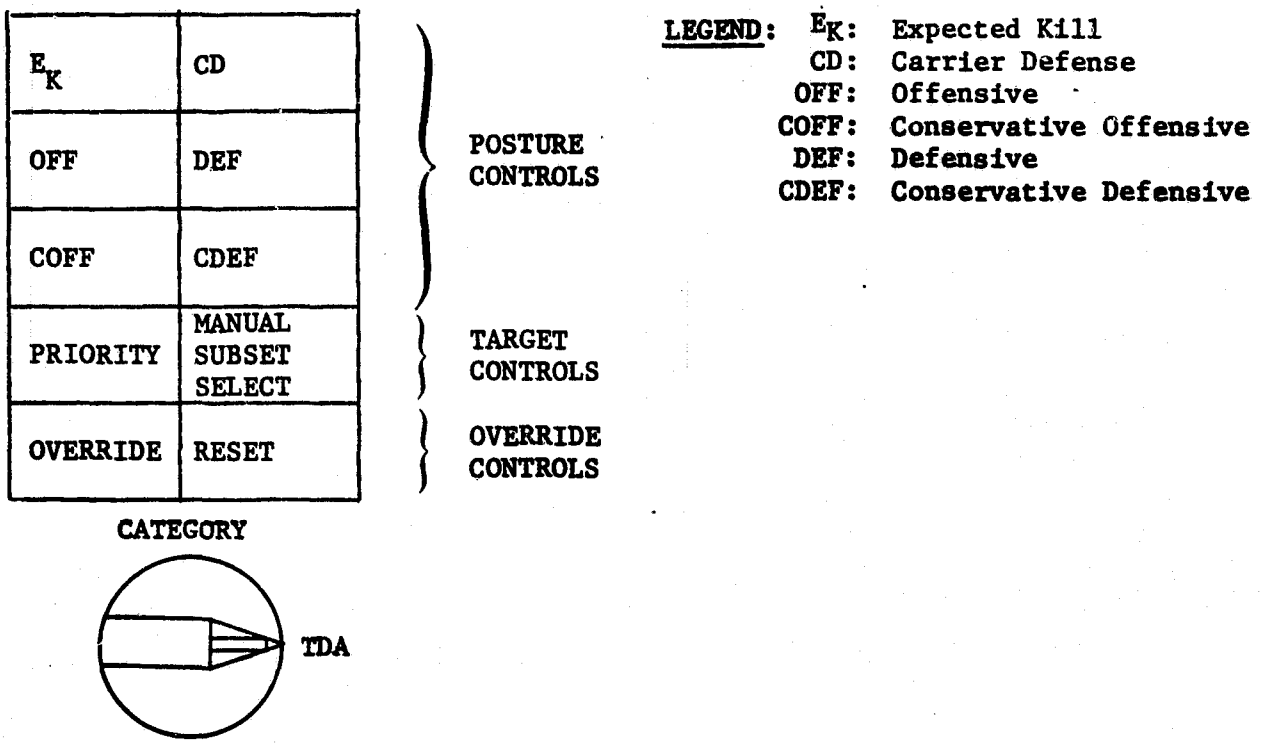


Figure 7. TDA Controls

Posture Selection Controls. The top six set of buttons correspond to the six postures that can be manually selected by the NFO. A button press results in the button being backlit signalling selection of the associated posture. This feature allows the NFO to keep track of the prevailing posture at all times. The postures are automatically selected by the transition program if no NFO posture-selection is indicated. If the NFO presses a posture button to select the posture of his choice, then this selection disengages the TDA posture transition logic and his selection posture will then be "frozen" regardless of conditions (and displayed just below the TID buffer readouts) until the NFO decides to reset the automatic posture selection logic of the TDA.

Target Selection Controls. The target selection section of the panel consists of two buttons: (a) priority target, (b) manual subset select. The "priority" target button allows the NFO to manually hook one target on the TID. This target is then included in all target subsets considered for subsequent evaluation. Priority target selection is accomplished by first pressing the priority target button and then hooking the target on the TID. The "priority" target button responds to button press by flashing and continues to flash until the target is hooked. The "manual subset select" button allows the NFO to manually designate the entire subset. When first pressed, this button starts to flash. The NFO then hooks each target in turn, until all targets have been hooked. The NFO indicates completion of subset selection to the system by pressing the manual subset select button once again. When the button is pressed for the second time, the flashing light goes off.

Override Controls and Maneuver Displays. The two basic override controls are available in the form of "reset" and "override" buttons, shown in Figure 7. The "reset" button allows the NFO to re-engage the automatic posture transition logic portion of the TDA. Once this software is invoked, target subset recommendation consistent with the prevailing posture is made automatically by the TDA and displayed on the TID along with the necessary course perturbation, if any, being indicated by a shift of the steering dot on the TID. There are at least three distinctively different occasions when the NFO might wish to use the "reset" button: (1) after selecting a posture manually, he decides against the selection; (2) after pressing the priority target button, he decides he does not want to hook a priority target or hooks a target and decides against it; (3) after pressing manual subset select he decides against it or after hooking a subset of targets he decides that one or more of them are inappropriate. For case (2) or (3), if the posture was selected manually, pressing reset puts the TDA in automatic posture select. If the NFO prefers to stay with his original manual selection of posture, he must reselect the posture manually after pressing "reset." The "reset" button requires two presses. The reason for this is that if accidentally bumped once it will not "erase" the target subset selected thus far by the NFO and/or deselect his manual posture selection. Consequently, when pressed once, the "reset" button causes a warning to be displayed on the DD, when pressed a second time it disengages the targets selected thus far and turns off the priority or manual subset select button backlighting. It also deselects the posture if it were manually selected by the NFO and turns off the backlit posture button. The "override" button allows the NFO to reject (1) a recommended aspect maneuver and/or (2) a recommended perturbation maneuver with associated subset.

When an aspect maneuver is recommended, the steering dot is moved on the TID and the aspect recommended is displayed as "ASPECT MAN" just below the TID buffer readout. When the override button is pressed following an aspect recommendation, the steering dot is returned to its previous position, the display area assigned to display "ASPECT MAN" is erased, and the TDA program branches immediately to subset select rather than waiting for the aspect maneuver to be performed. The other recommendation provided by the TDA is that of optimum target subset selection with associated perturbation. When this recommendation occurs, the subset is brightened, the steering dot is moved on the TID and "LAR ACQ MAN" is displayed just below the TID buffer readout. If "override" is pressed during a subset recommendation, the "next best" subset with associated perturbation is presented, the TID steering dot is moved, "LAR ACQ MAN" is displayed as before, and subset display on the TID modified accordingly. If reject is exercised again, the "LAR ACQ MAN" display is erased but the "next best" subset remains the same (i.e., does not change from the previous display). In this case, since the NFO finds the TDA's recommendation unsatisfactory, he can select the targets manually by "hooking" them on the TID after pressing manual subset select.

TDA Override and Restart Options

There are three TDA-related KIWI activities that were found to be logical abort points for establishing new parameters and re-directing TDA processing. These are:

(1) Aspect maneuver.
(2) Perturbation maneuver.
(3) Steering to encounter.

Aspect Maneuver Override. The "override" button may be pressed by the NFO at any time during the aspect maneuver sequence. When this happens, only the aspect maneuver is aborted, thus, as far as the decision logic is concerned, the aspect maneuver is assumed to be completed and the subsequent TDA activities remain unaltered.

Perturbation Maneuver Override. The "override" button may once again be pressed by the NFO anytime during the perturbation maneuver. This action causes control to be immediately passed to the steering algorithm rather than wait for completion of the perturbation maneuver before passing control to the steering algorithm.

The restart options that can be exercised during any of the three logical abort points include manual posture selection, target selection, and TDA reset

Manual Posture Selection. Selecting a posture manually by pressing the intended key on the CAP panel will discontinue TDA or KIWI activity at any of the three abort points and pass control to the subset optimization procedure. If posture is altered prior to target detection, the TDA decision sequence is unaffected.

Target Selection. Pressing "manual subset select" or "priority target" disengages the TDA or KIWI at any of the three abort points. After manually selecting targets with "manual subset select," control is passed immediately to the steering algorithm. When a priority target is selected, control is passed to the subset optimization procedure in the TDA. Optimization is performed with the constraint that the priority target must be contained in the "optimum subset."

TDA Reset. When this button is pressed, the TDA or KIWI is disabled at any of the three abort points and control is passed to the initial target acquisition phase of KIWI with automatic posture updating.

CONCLUDING REMARKS

This paper has presented a realtime Tactical Decision Aid (TDA) for the F-14 aircrew in air intercept operations associated with the Combat Air Patrol (CAP) role. The aid has been designed with special emphasis on ensuring that (1) its operation fits naturally within the task structure of the aircrew; and (2) it in no way appears to usurp any of the aircrews' traditional activities. While it was recognized that there is strong preference among the user community for the display of options rather than the optimum (Mackie, 1980), the short itme horizons associated with air intercept tactics preclude the aircrew from scanning and cogitating the various alternatives. Consequently, the aid allows the aircrew to exercise manual override over any recommendation it makes. The aid is currently being implemented in the multiple target environment of the KIWI simulator, a realtime AWG-9 simulation at Naval Air Development Center, Warminster, Pennsylvania.

REFERENCES

Dawes, R. M. and Corrigan, B. Linear Models in Decision Making. Psychological Bulletin, 1974.

Development and Application of a Decision Aid for Tactical Control of Battlefield Operations. Volume 1, A Conceptual Structure for Decision Support in Tactical Operations Systems, Honeywell Systems & Research Center, August 1974.

Fischer, G. W. Multi-Dimensional Value Assessment for Decision Making. Engineering Psychology Lab., University of Michigan, Tech. Report 037230-2-T, 1972.

Gardiner, P. C. The Application of Decision Technology in Monte Carlo Simulation to Multiple Objective Public Policy Decision Making: A Case Study in California Coastal Zone Management. Unpublished Ph.D. dissertation, Center of Urban Affairs, University of California, 1974.

Madni, A., Steeb, R., and Freedy, A. Computer-Based Tactical Decision Aids for the F-14 AWG-9 Weapon System, Perceptronics Final Technical Report P-79-102-1079, May 1979.

Madni, A., Steeb, R., and Purcell, D, F-14 Tactical Decision Aid: System Description and KIWI Implementation. Perceptronics Final Technical Report PDFTR-1086-80-6, May 1980.

Mackie, R. R. Design Criteria for Decision Aids: The Users Perspective, Proceedings of the Human Factors Society-24th Annual Meeting, 1980.

Meador, C. L. and Ness, D. N. A case study of the use of a decision support system. Massachusetts Institute of Technology, 674-73, 1973.

Morton, M. S. S. Decision Support Systems: The Design Process. Sloan School of Management Report, No. 686-73, November 1973.

ACKNOWLEDGEMENT

This work was sponsored by Naval Air Development Center, Warminster, Pennsylvania under Contract No. N62269-79-C-0496. The Technical Monitor for the project was Dr. Mark Elfont. The authors wish to acknowledge the contributions of Dr. Randall Steeb and Mr. Denis Purcell who assisted in problem structuring modeling, and implementation. Appreciation is extended to the F-14 Naval Flight Officers and pilots at Veda, San Diego who provided the subject matter expertise.

USING REWARDS AND PENALTIES TO OBTAIN DESIRED SUBJECT PERFORMANCE

Marcia Cook, Henry R. Jex,
Anthony C. Stein and R. Wade Allen
Systems Technology, Inc.
Hawthorne, California 90250

SUMMARY

The Critical Tracking Test (CTT) is a psychomotor test that has proven in the past to allow reliable measurement of human operator performance, and has been shown to give sensitive decrements to a variety of stresses. Currently the CTT is being tested as a method of detecting human operator impairment. A pass level is set for each subject, based on that subject's asymptotic skill level while sober. It is critical that complete training take place before the individualized pass level is set in order that the impairment can be detected.

Some subjects have shown erratic learning trends typical of unmotivated performance. These subjects seemed more motivated to finish their training sessions quickly, rather than to receive monetary rewards for good performance.

This paper describes operant conditioning procedures, specifically the use of negative reinforcement, in achieving stable learning behavior. These results now provide a more general basis for the application of reward/penalty structures in manual control research.

INTRODUCTION

The Critical Tracking Test (CTT) is a psychomotor task based on manual control principles (Reference 1) that has been highly developed and applied to a variety of research problems (Reference 2). CTT performance is affected by individual differences and abilities, and critical instability scores (λ_c) vary across subjects. For a given trained subject, however, λ_c is quite consistent, and is sensitive to blood alcohol concentration. Thus, it is possible to set a pass/fail level required for a particular decision strategy (Reference 2). It is important that the pass criterion accurately reflect the subject's skill level. If λ_{pass} is too low, the operator will be able to pass while intoxicated. If λ_{pass} is too high, the sober operator will experience unnecessary and annoying delays while the equipment resets. One objective of this experiment was the development and refinement of a rational a priori

procedure for setting pass levels based on training scores only. Because of motivational problems encountered in the original Experiment, a subsequent Training Experiment was carried out.

EXPERIMENT

Background

For the experiment, 24 subjects were obtained through the Los Angeles Municipal Courts, who had plead guilty to charges of Driving Under the Influence of Alcohol (DUI). The judge offered them the option of participating in this research project instead of the usual $350 traffic fine and traffic school. Even though the monetary rewards for participating were substantial, it was still a choice between the lesser of two evils. Subject participation was often reluctant, and in some cases subjects were effectively participating under duress.

The MMPI was administered and subjects with clinically abnormal profiles were eliminated from the population. The subjects were trained to "drive" in a simulator and to perform the critical tracking task. Three sessions of approximately 2 hours each on separate days were required for training. Following training, subjects participated in 3 experimental alcohol sessions, scheduled about a week apart, and lasting 10 to 12 hours. Each subject also participated in a follow-up session of approximately 1/2 hour. They received $3.10/hour for training and experimental sessions, and a bonus schedule was worked out to provide incentives for good performance and completion of the program (e.g., Reference 2). The subjects averaged about $220 total from wages and bonus money and the court cancelled a $350 traffic fine if they satisfactorily completed all requirements.

Learning curves were generated from training data to determine if asymptotic levels of performance had been attained. The 24 training plots were evaluated for indications of poor motivation or inadequate training. Half of the subjects showed lack of motivation and/or not enough training. These curves were characterized by unusually low scores in the last training session and a slope that was increasing even at the end of the last session (further discussion of this problem is found in section entitled Results). The remaining subjects showed asymptotic learning and consistent behavior, i.e., Figure 1. The simple pass criteria chosen for these 12 subjects correlated 0.862 with the "ideal" or hindsight criteria taken from the later experimental data as shown in Figure 2.

Methods

CTT training was done in 3 sessions on separate days. Each subject read the instructions (Appendix) and was informed of the bonus structure (at least 1 pass in a block of 4 trials = 75¢). Each training session consisted of 25 blocks of 4 trials, or 100 trials.

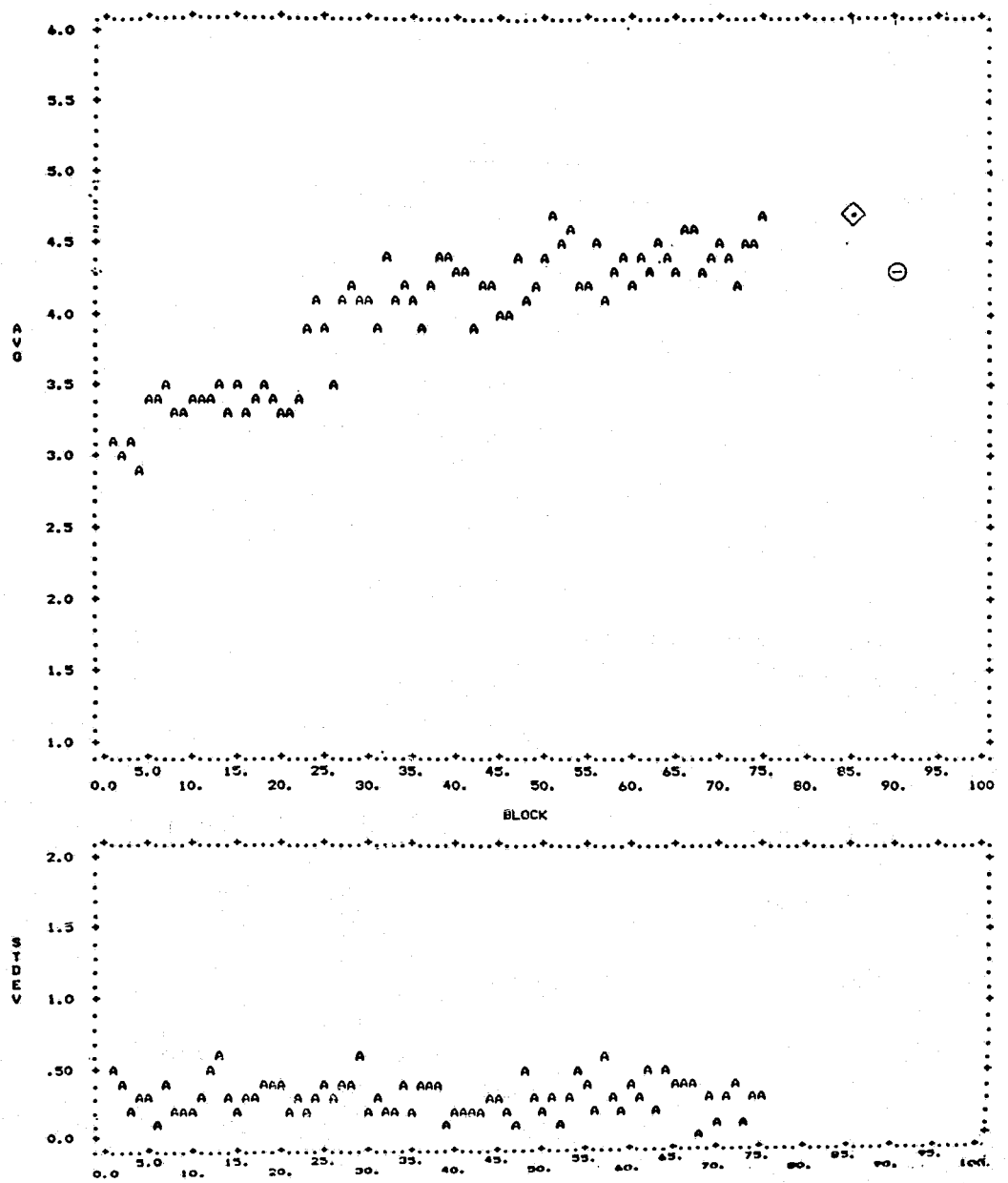


Figure 1. Good Asymptotic Training Exhibited by One Test Subject

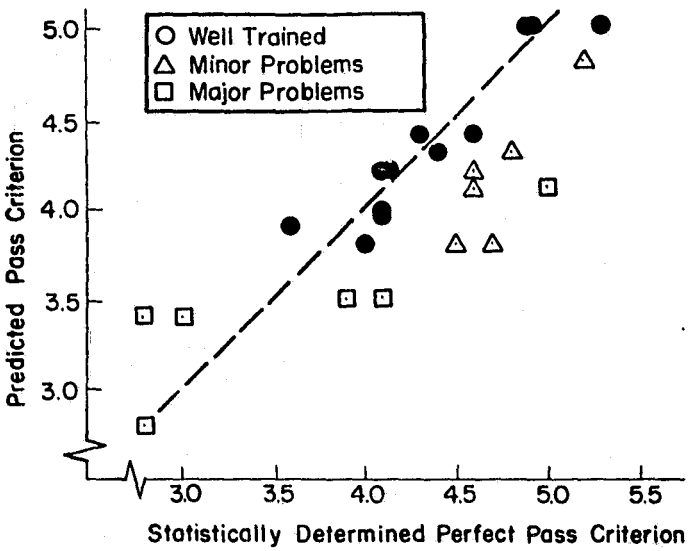


Figure 2. Pass Criterion Analysis

The initial pass criterion was set at an arbitrary low level that all subjects could pass within a few trials (i.e., $\lambda = 2.9$). If the subject passed 3 out of 4 trials in a block, the pass level was raised 0.1. If the subject passed all 4 trials the pass level was raised 0.2. If the subject failed all 4 trials the pass level was lowered 0.1. The pass level could go back to, but not below, a level where 3 out of 4 trials had previously been passed. This strategy tends to prevent deliberate "backsliding" by the subjects.

Breaks were given every 10 or 15 minutes. The mean λ_c for each block in the last training session was computed. The experimental λ_p was chosen by taking the median of the highest 3 block means and reducing this value by 0.3 (roughly 1/2 of σ_λ). This procedure was found to give a pass level consistent with a one trial probability of failure of 40 percent which was required for a one-pass-in-four-attempts decision strategy as discussed in Reference 2.

Results

For this experiment we were interested in determining the extent to which we were able to predict failure rates of impaired operators based on sober training data, and thus needed to predict an accurate pass score (λ_{pass}). The experimental (alcohol) session data (all sober baselines and placebo trials) were analyzed to determine in hindsight what the "ideal" (a posteriori) pass/fail score for each subject should have been to insure a 40 percent probability of failing a single trial when sober (as discussed in Reference 2). Comparing the experimental criterion set during this ideal criterion we found that in 15 cases the

λ_{pass} was too low (Figure 2). The solid circles indicate those subjects who were well trained. For these individuals our predicted pass level was remarkably consistent with the statistically determined perfect pass level. Problems were found when we compared predicted and perfect pass levels for subjects with both minor and major training problems (i.e., erratic learning curves). Subjects with minor problems exhibited a remarkable and consistent increase in λ_c scores as soon as there was no time benefit to failing. Subjects with major training problems had displayed problems consistently during training, and these problems persisted in the experiments. Even though the individual pass criteria were not perfect, the discriminability results agree with the statistical model described in the companion paper (Reference 2).

Discussion

Because 15 of the 24 subjects were assigned λ_{pass} criteria that were too low, the incentive structure was re-evaluated. The 75¢ bonus had been given for passing at least one trial in a block of 4. (This was chosen because it simulates the one pass out of four attempts strategy used in the formal validation experiment trials in the reanalysis of prior data (Reference 2). We found that most of the subjects did not even bother to learn the reward structure in order to maximize the bonuses earned (the bonus money was added across sessions and paid out at the completion of the experiment). They preferred to put forth a random effort and accept whatever total monies they happened to earn.

Research in operant conditioning (i.e., Reference 4) has shown that the more delayed such a "reinforcer" is, the less potent it is. Therefore, while the bonus money may have reinforced the subject for completing the experiment, it had little effect on each individual trial of the CTT. It was also observed during training that the subjects' primary motivation was to complete training sessions as soon as possible, even to the extent of foregoing rest breaks and failing the test quickly to speed up the trial repetition rate.

It was decided that in order to elicit a consistent and stable performance from the subject a more immediate reinforcer should be applied after each passed trial on the CTT. Reinforcement occurs when a rewarding stimulus follows a response or when a negative or punishing stimulus is avoided. It was decided that, since the training procedure is so tedious, a negative, or punishing, stimulus would be a 30 second "time out" condition added after each failed trial. By using the avoidance paradigm, the absence of the aversive stimulus (the "time out") becomes a reinforcer (Reference 4).

A subsequent study was conducted to verify if addition of this "time out" procedure would motivate the subjects in order that stable and complete training is obtained in three sessions.

TRAINING EXPERIMENT

Objectives

As described earlier, each subject is required to complete 300 trials on the CTT to establish an individualized pass/fail score. This is a tedious process done over 3 sessions, each lasting 2 to 3 hours.

Because of the poor performance and comments made during the experiment, we conducted a brief training experiment using the reinforcement strategy described above. In this situation a green "pass" light in the display of the CTT apparatus takes on new reinforcing properties, as it now signals the absence of a "time out." In the previous experiment the green light, in general, only provided information as to the outcome of the trial. The red "fail" light, on the other hand, now takes on aversive properties because it is present during the 30 second "time out." The red light was also merely informational in the previous experiment.

Procedure

Six additional subjects were contacted and enrolled through the Los Angeles Courts, as before. One hundred dollars of their $350 fine was dropped when they completed the project. The MMPI was administered as in the earlier Experiment. The first two subjects were given the printed instructions (Appendix) that outlined the objectives of the study and the bonus structure. They repeatedly expressed their surprise at being paid, because traffic schools do not pay for participation. Also, as in the earlier experiment, they chose not to learn the incentive structure and said, in effect, "You just keep track and pay me later." It was decided at that point that paying the subjects was superfluous and unnecessary, so the next four subjects received verbal instructions with no mention of wages or incentives.

The subjects received the same training procedure as in the earlier experiment (25 blocks of 4 trials each), except that each time a trial was failed there was a 30 second delay. [Note: These subjects all came after work and, because of the hour and the time of year, the testing, which took place in a parked car, was done in the dark. When the trial was completed, the display light went off with only the red "FAIL" or green "PASS" light remaining on, depending on the outcome of the trial. This meant that for a failed trial the subject waited in a dark car for 30 seconds, looking at a small red light that said "FAIL."]

Results

Some of the learning curves are shown in Figures 3-5 for the new condition using the "time out" procedure. There is no qualitative difference between those subjects that were paid and those that were not paid. The learning curves show the same asymptotic learning curves that

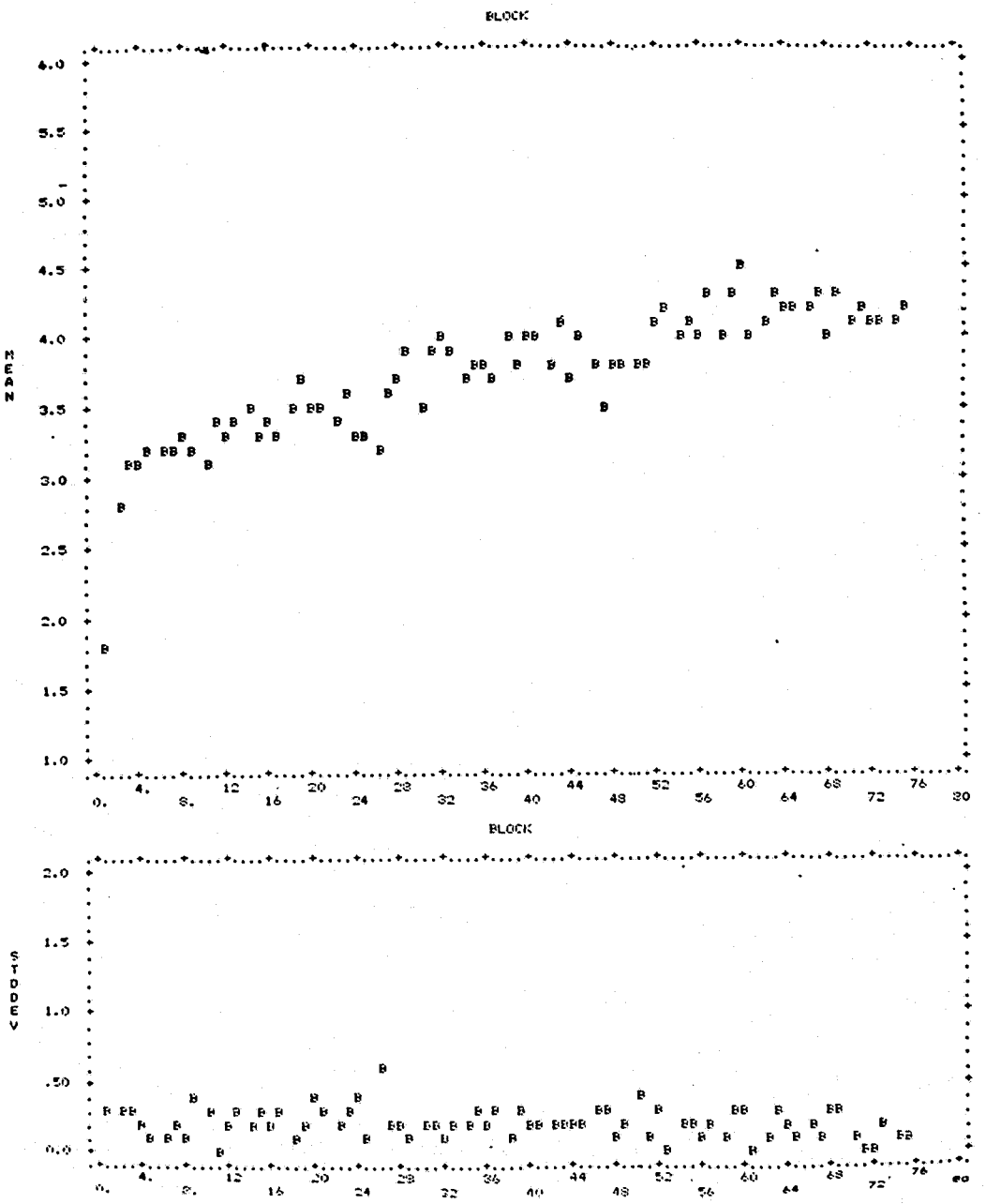


Figure 3. Learning Curve for a Subject Who Received Delayed Monetary Rewards for Task Performance in Addition to Immediate Wait Time Penalties

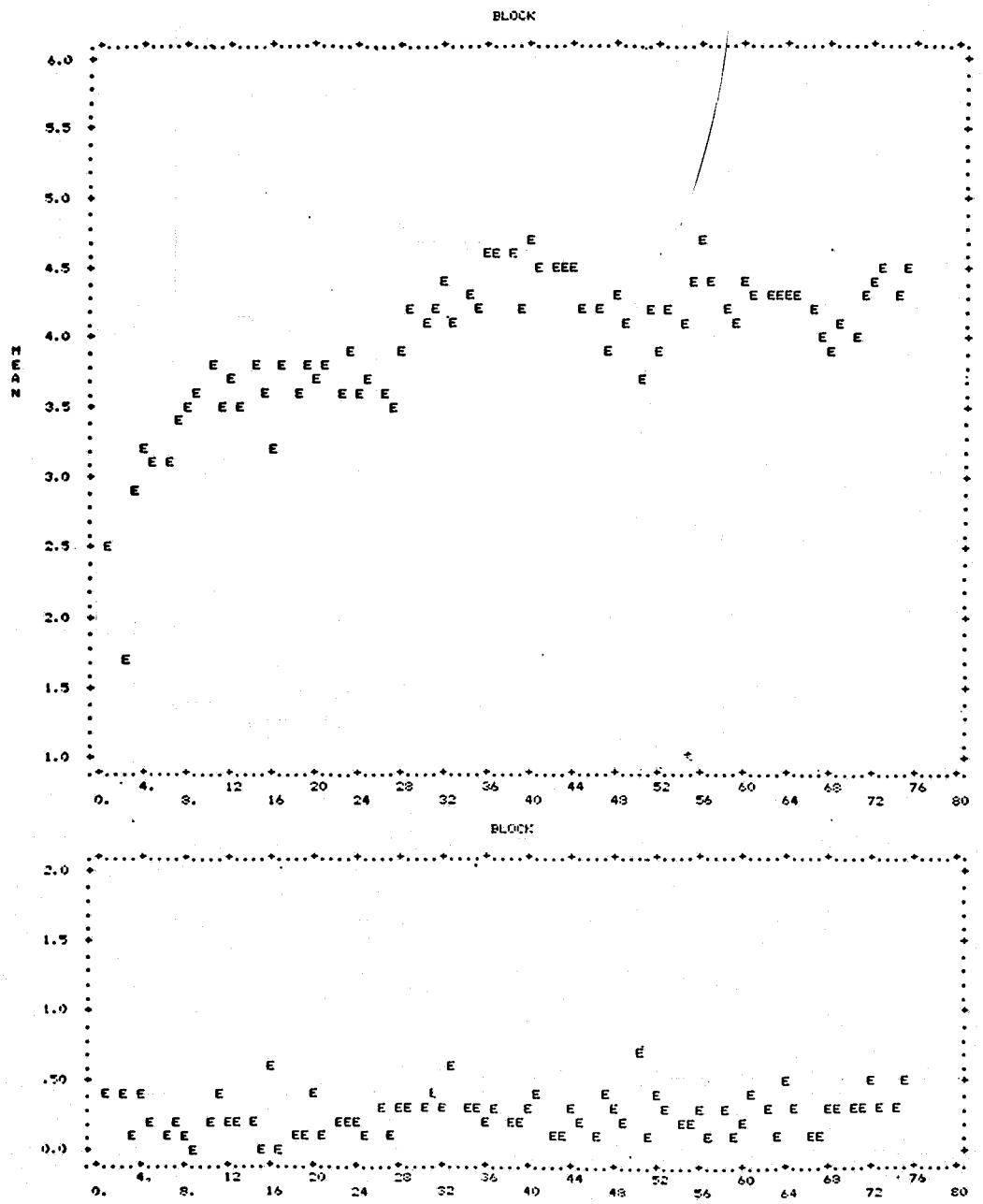


Figure 4. Learning Curve for a Subject Receiving No Monetary Rewards for Task Performance

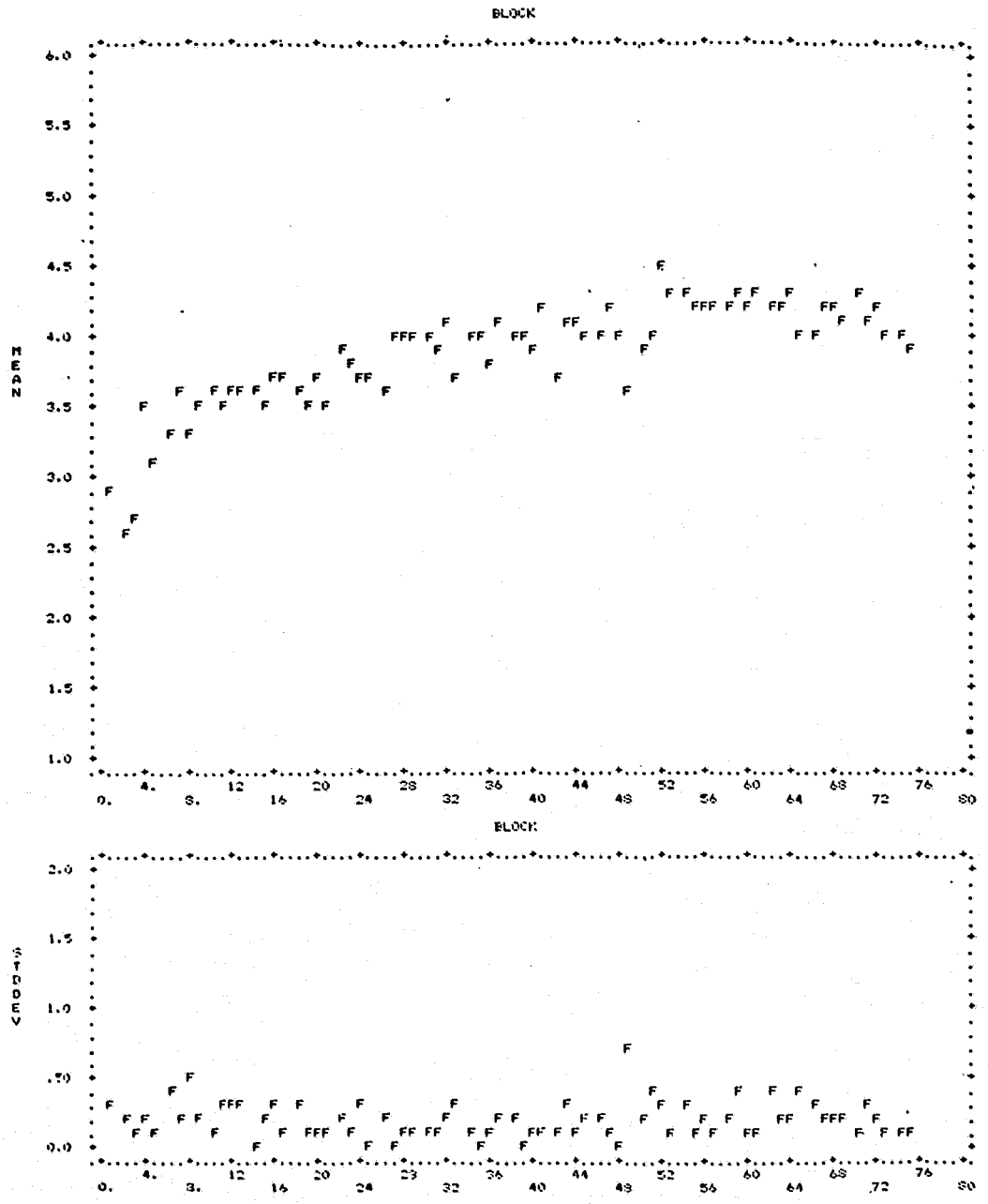


Figure 5. Learning Curve for a Second Subject Receiving No Monetary Rewards for Task Performance

yielded accurate criteria predictions in the earlier experiment. Furthermore, the trial-to-trial consistency is better, as shown by the low standard deviations at the bottom of Figures 3-5.

CONCLUDING REMARKS

This study has shown that, to make an accurate prediction of a pass criterion, stable and complete learning must take place. It has also been shown that learning is facilitated when the subject is reinforced after each trial. In the past, subjects were given numerical feedback on each CTT trial as to what their actual scores were (i.e., References 5-7). This acted as an immediate and powerful reinforcer, because each trial provided an opportunity to better their last score. In addition, we were fortunate in working only with motivated subjects. For these reasons our earlier Reward/Penalty structures (Reference 3) were based on a positive reinforcement model only. In this experiment the subject was trained to "pass the test" without being aware that the eventual pass level was adjusted to their individual ability to perform. Since the normal reinforcement by display of CTT scores was not used, the 30 second time out after each fail was substituted.

In assessing what reinforcers are available for use with unmotivated subjects, money seems to be a neutral stimulus. The main motivator should be based on passing the test to avoid the aversive stimulus. In this way subjects will learn to pass the test without being aware that the eventual pass level is adjusted according to their individual ability to perform. The "30 second time out for fail" procedure seems to provide the desired immediate reinforcement after each trial and maintains consistent test scores. In addition, these results allow for a more general application of reward/penalty structures in manual control research depending on the basic motivation for subject participation.

REFERENCES

1. Jex, H. R., J. D. McDonnell, and A. V. Phatak, A "Critical" Tracking for Man-Machine Research Related to the Operator's Effective Delay Time. Part 1: Theory and Experiments with a First-Order Divergent Controlled Element, NASA CR-616, Nov. 1966.

2. Allen, R. Wade, Anthony C. Stein, and Henry R. Jex, "Detecting Human Operator Impairment with a Psychomotor Task," 1981 Annual Conference on Manual Control Proceedings, Jun. 1981, in press.

3. Stein, Anthony C., R. Wade Allen, and Stephen H. Schwartz, "Use of Reward-Penalty Structures in Human Experimentation," 14th Annual Conference on Manual Control, NASA CP-2060, Nov. 1978, pp. 267-278.

4. Reynolds, George S., A Primer of Operant Conditioning, Glenview, IL, Scott, Foresman, 1968.

5. Oates, John F., Jr., Experimental Evaluation of Second-Generation Alcohol Safety-Interlock Systems, DOT-TSC-NHTSA-73-9, 1973.

6. Oates, John F., David F. Preusser, Richard D. Blomberg, and Marlene S. Orban, Laboratory Testing of Alcohol Safety Interlock Systems. Vol. I: Procedures and Preliminary Analyses, Dunlap and Associates, Darien, CN, Jun. 1975.

7. Oates, John F., David F. Preusser, and Richard D. Blomberg, Laboratory Testing of Alcohol Safety Interlock Systems, Phase II, Dunlap and Associates, Darien, CN, Dec. 1975.

APPENDIX

This study is being conducted for the U.S. Department of Transportation to test a device designed to tell a driver that he may be too drunk to drive. Your participation in this project is important because you will be contributing to the improvement of measures for reducing alcohol related accidents. This phase of the study is aimed at helping us finalize our training strategy for the task.

You will use the device to perform a steering control task. We will provide you with an opportunity to become familiar with the operation of the device during three separate sessions. Each session will last approximately three hours.

The steering control task will be performed on a unit which is mounted on the side of the steering wheel in a specially modified car. The device consists of a meter (a display and a needle), a start switch, and a small computer in the trunk of the car. When the task begins, the needle is centered in the green area of the display. As the task proceeds, the needle begins to wander either to the left or the right. It is your job to keep the needle centered in the green area by moving the steering wheel in the direction you want the needle to move. It will become increasingly difficult to keep the needle centered.

The device will automatically set a passing level for your performance. This pass level is based on learning rates established in previous research. Trials will be conducted in groups of four. If you pass one trial in a group of four, you will receive a cash reward. If you have already passed one trial, it is important to continue to do your best on the remaining trials, since the task will continue to increase in difficulty, and your ability to pass future trials will be affected. The level of difficulty will change after each set of four trials. It will be necessay for the experimenter to reset the system periodically.

We expect that you will do poorly on the first few trials because you are completely unfamiliar with the task. You should not become discouraged; we know that your performance will improve with training.

You will be paid $3/hour plus a bonus of 75 cents per group passed for each session. You can expect to average $12-14 per session/

Your pay will be divided up into 2 parts — each day you will take home your hourly wages, thus on an average day you take home about $9.00. Your bonus money is held in a holding account, to be paid to you on your last day of experiments (approximately $13.50 in addition to your hourly pay).

Should you drop out sometime during the experiment, your holding account money will still be owed to you, but it can't be collected until the completion of the experiments, and you will have to collect in person.

D20

A COMPARISON OF LANDING MANEUVER PILOTING TECHNIQUE BASED ON MEASUREMENTS MADE IN AN AIRLINE TRAINING SIMULATOR AND IN ACTUAL FLIGHT

Robert K. Heffley and Ted M. Schulman

Systems Technology, Inc.

SUMMARY

An analysis of pilot behavior, both from an airline training simulator and an actual DC-10, is presented for the landing maneuver. An emphasis is placed on developing a mathematical model in order to identify useful metrics, quantify piloting technique, and define simulator fidelity. On the basis of DC-10 flight measurements recorded for 32 pilots — 13 flight-trained and the remainder simulator-trained — a revised model of the landing flare is hypothesized which accounts for reduction of sink rate and preference for touchdown point along the runway. The flare maneuver and touchdown point adjustment can be described by a pitch attitude command pilot guidance law consisting of altitude and vertical velocity feedbacks. The pilot gains which are identified directly from the flight and simulator data show that the flare is being executed differently in each medium. In flight most of the subject pilots exhibit a significant vertical velocity feedback which is essential for well controlled sink rate reduction at the desired level of response (bandwidth). In the simulator, however, the vertical velocity feedback appears ineffectual and leads to substantially inferior landing performance. The absence of the vertical velocity feedback implies a simulator fidelity problem, and several specific possibilities are discussed. The pilot model of the maneuver provides insight into which aircraft types could be simulated without incurring the apparent fidelity limitation encountered in this case.

INTRODUCTION

This paper is a summary of portions of an analysis of airline landing data which was performed for NASA Ames Research Center under Contract NAS2-10817 and reported in Ref. 1. The purpose of the study was to focus on the landing maneuver as it is performed both in flight and in an airline training simulator in order to: (a) measure absolute differences between pilot-vehicle behavior, (b) develop landing maneuver performance metrics, and (c) define how to use such metrics in both simulator and flight.

The data base used in this analysis was collected during a NASA field evaluation of the sole use of simulator training in transitioning airline

pilots to a new aircraft type (Ref. 2). The unique aspect of the data acquired is that they involve both actual flight and simulator measurements for a reasonably large number of pilots. Furthermore specific attention was devoted to making the flight and simulator data directly comparable in terms of pilots, aircraft, and environmental conditions.

The procedure used in analyzing the available data was based on manual control theory (Ref. 3) which treats human psychomotor and cognitive behavior as rational, well-tailored actions dependent upon the task, vehicle dynamics, and environment. These actions can be essentially closed loop and compensatory in nature or progressively more open loop and precognitive depending upon the pilot's level of skill or workload demands.

The issue of simulator fidelity has been stated in terms of manual control theory in Ref. 4 and is highly relevant to the analysis. In fact perceptual fidelity is addressed in terms of "essential cueing" as discussed in Ref. 5. As will be seen, there is evidence that the training simulator involved in this study was somehow deficient in inducing the pilot behavior observed in flight. This kind of deficiency should be duly noted in the design and actual use of any simulator where flight task and aircraft conditions are similar to those studied here.

SYMBOLS

h	Height
$\dot{h}_{TD}$	Touchdown sink rate
$\dot{h}_{max}$	Maximum sink rate
k_h	Pilot height loop gain
$k_{\dot{h}}$	Pilot vertical velocity loop gain
k_γ	Pilot flight path angle loop gain = $Uk_{\dot{h}}$
s	Laplace operator
T_{AC}	Effective aircraft flight path lag
U	Airspeed
ζ_{FL}	Effective damping ratio of the landing maneuver
θ_c	Pitch attitude command
ω_{FL}	Effective natural frequency of the landing maneuver
ω_{c_θ}	Pitch loop crossover frequency

FLARE MODEL

The appendix of Ref. 1 reviewed some existing models of the flare maneuver (Refs. 6 through 9), considering their strong and weak points. These ideas were taken into account in constructing a revised flare model which would better explain the recently-acquired landing data as well as encompassing past measurements. One important aspect of this revised model is that there is no added complexity over previous models discussed,

in fact there is significant reduction in complexity —— so much so that a closed analytic form can be expressed for time histories of altitude, sink rate, normal acceleration, airspeed decay, and touchdown point along the runway. Furthermore it is possible to describe a clear role for the important aircraft properties as well as for the pilot control law properties. This ultimately aids in developing metrics for analyzing the landing maneuver.

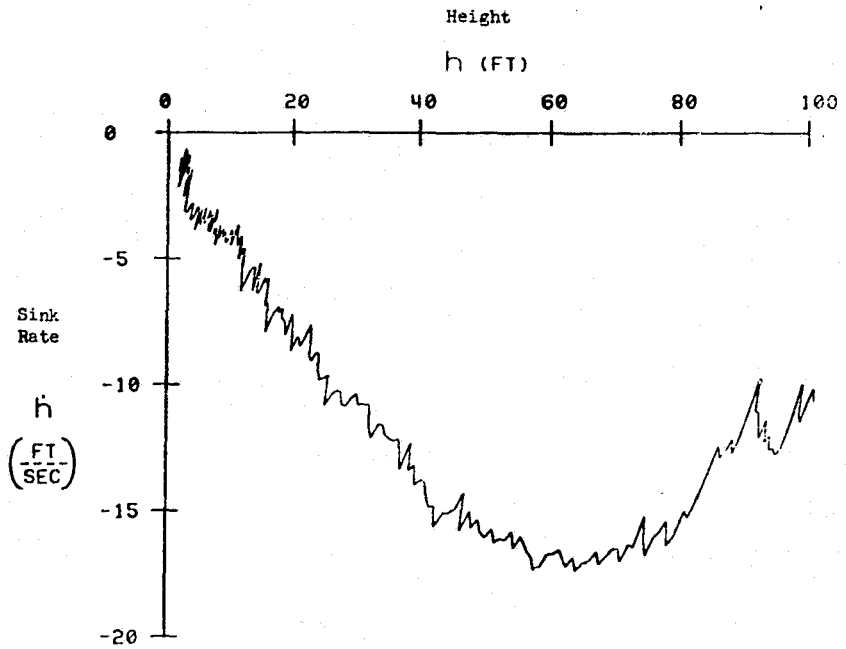


Figure 1. Phase Plane Trajectory of a Typical Landing Maneuver

The theoretical basis for the revised model is the assumption of dominant second-order characteristic response which is strongly suggested by the phase planes constructed from flight data of which Fig. 1 is an example. This leads to the basic characteristic equation:

$$\ddot{h} + 2\zeta_{FL}\omega_{FL}\dot{h} + \omega_{FL}^2 h = 0$$

It is further assumed that this characteristic equation is associated with a pilot-vehicle system having an altitude command loop (outer loop) and that the flare maneuver corresponds to the response from an initial offset with respect to the terminal conditions (i.e., from an initial altitude and sink rate). Thus, analytically, the flare is regarded as an unforced response from a set of initial conditions to a set of desired conditions at touchdown.

In considering the pilot control law implications of a second-order characteristic response, the first step is to examine the aircraft equations of motion with respect to altitude. The complete longitudinal formulation described in Ref. 10 can be simplified to a first-order, single-axis form:

$$\ddot{h} + \frac{1}{T_{AC}}\dot{h} = \frac{U}{T_{AC}}\theta_c$$

Where T_{AC} is the effective first-order lag time constant between a pitch-command, θ_c, and flight path response.

The approach used to infer piloting technique in the landing maneuver was to solve directly for the difference between a fitted differential equation describing closed-loop motion and the effective flight path response of the basic airplane. The difference, assuming negligible atmospheric disturbances, should be the effect of pilot actions and can be interpreted literally as a pilot control law, i.e.,

$$\ddot{h} + 2\zeta_{FL}\omega_{FL}\dot{h} + \omega_{FL}^2 h = 0 \qquad \text{(fitted differential equation of landing maneuver)}$$

minus $$\ddot{h} + \frac{1}{T_{AC}}\dot{h} = \frac{U}{T_{AC}}\theta_c \qquad \text{(Aircraft flight path equation)}$$

---

equals $$\left(2\zeta_{FL}\omega_{FL} - \frac{1}{T_{AC}}\right)\dot{h} + \omega_{FL}^2 h = -\frac{U}{T_{AC}}\theta_c \qquad \text{(inferred pilot control law)}$$

Rearranging the result,

$$\theta_c = \underbrace{-\frac{\omega_{FL}^2 T_{AC}}{U}}_{k_h} h - \underbrace{\frac{2\zeta_{FL}\omega_{FL}T_{AC} - 1}{U}}_{k_{\dot{h}}} \dot{h}$$

Hence the effective control law gains can be related directly to parameters describing the maneuver and the aircraft:

$$k_h = \frac{\omega_{FL}^2 T_{AC}}{U}$$

$$k_{\dot{h}} = \frac{2\zeta_{FL}\omega_{FL}T_{AC} - 1}{U}$$

or basing a control law term on flight path angle, γ, rather than sink rate, $\dot{h}$:

$$k_\gamma = 2\zeta_{FL}\omega_{FL}T_{AC} - 1$$

IDENTIFICATION OF PARAMETERS

The foregoing theoretical development shows that two motion parameters and one aircraft parameter are needed to obtain the effective pilot control law parameters.

The aircraft parameter, T_{AC}, was obtained from an estimate of the flight path response for a pitch attitude command in a DC-10 at an average landing weight and speed. Allowance was made for the contribution of lag in the closed-loop pitch response as well as the lag due to airframe heave damping, T_{θ_2}. An effective first-order pitch response lag of 0.7 sec was assumed based on previously observed transport aircraft pitch attitude closures ranging from 1 to 2 rad/sec crossover frequency, ω_{c_θ}. For the airframe heave damping component, a value of 1.8 sec was estimated for the average T_{θ_2} corresponding to the loading and approach airspeeds flown. A composite flight path response lag, T_{AC}, was obtained by summing the pitch response and heave response lags, i.e.,

$$T_{AC} \triangleq T_{\theta_2} + \frac{1}{\omega_{c_\theta}} \cong 2.5 \text{ sec}$$

This approximation can be shown to be valid for landing maneuvers having an effective damping ratio, ζ_{FL}, in the vicinity of 0.7 —— the nominal value found in the flight data.

The landing maneuver was identified directly from phase plane trajectories plotted for the flight and simulator landings. Two separate procedures were developed for obtaining independently the effective flare damping ratio, ζ_{FL}, and the effective natural frequency, ω_{FL}. It was found that a strong relationship existed between ζ_{FL} and the ratio of touchdown sink rate to maximum sink rate (just prior to flare), $\dot{h}_{TD}/\dot{h}_{max}$.

The effective natural frequency of the flare, ω_{FL}, was shown to be a strong function of the shape of the flare trajectory and nearly independent of ζ_{FL}. As a consequence it was possible to identify ω_{FL} using transparent overlays of families of phase plane trajectories.

RESULTS OBTAINED

Nominal Landing Maneuver

The check-ride landings of the flight-trained group of pilots were used to obtain an indication of the nominal landing maneuver for the

DC-10. Figure 2 shows samples of the flight data in terms of landing trajectory phase planes for several pilots. Note that for each of the five pilots there were three landings. Figure 3 shows the identified landing maneuver parameters, ζ_{FL} and ω_{FL}, for these pilots along with a plot of pilot control law parameters. (Note that the height loop gain, k_h, is plotted opposite ω_{FL}^2 and the effective flight path angle gain, k_γ, is plotted opposite $\zeta_{FL}\omega_{FL}$.) Means and standard deviations of nominal landing parameters are summarized in Table 1.

It can be seen that the nominal flare maneuver parameters are grouped in rational locations with respect to the several factors, previously mentioned, which affect the landing; namely, the natural frequency —— an indicator of closed-loop bandwidth —— is situated midway between the closed-loop airspeed response mode (about 0.1 rad/sec for this aircraft) and the closed-loop pitch response (about 1 to 2 rad/sec). This partitioning of frequency helps to insure that airspeed will not bleed off excessively during the landing maneuver, and that pitch response will not significantly detract from the heave response phase margin. (If pitch loop crossover is set too close to flight path crossover, then a K/s^2-like controlled element is created.) The nominal value of damping ratio centered at about 0.7 helps to insure that a good touchdown sink rate is obtained regardless of the conditions at flare initiation. Too low a damping ratio, say 0.4, would correspond to a hard landing even from a nominal approach sink rate. At the other extreme, a damping ratio greater than, say, 0.9 would correspond to a floating tendency resulting in excessive runway landing distance. The nominal $\zeta_{FL} \cong 0.7$ and $\omega_{FL} \cong 0.4$ rad/sec are therefore entirely appropriate from the standpoint of good closed-loop control considerations.

The nominal piloting technique parameters spanned a range of effective loop gains. Most noteworthy, however, is that some degree of sink-rate or flight-path-angle feedback is apparent except where a very low height gain is employed. For the average k_h of 0.13 deg/ft*, an average k_γ of 0.45 (or $k_{\dot{h}}$ of 0.12 deg/ft/sec) was observed. This in turn implies that cues in addition to height may be used by the pilot. These data, however, did not indicate which of the several visual (or even motion) cues might have been involved.

In addition to the two landing maneuver parameters, ζ_{FL} and ω_{FL}, attention was given to how to characterize initiation of the landing. Flare height has been generally regarded as a likely candidate for a landing parameter, but the data showed no clear tendencies. Instead a wide range of heights for flare initiation were observed. Also most landings involved a "duck-under," i.e., an increase in sink rate, just prior to the

---

* This agrees with the DC-10 flight manual procedure —— about 3.5 deg net pitch change over the final 30 to 40 ft.

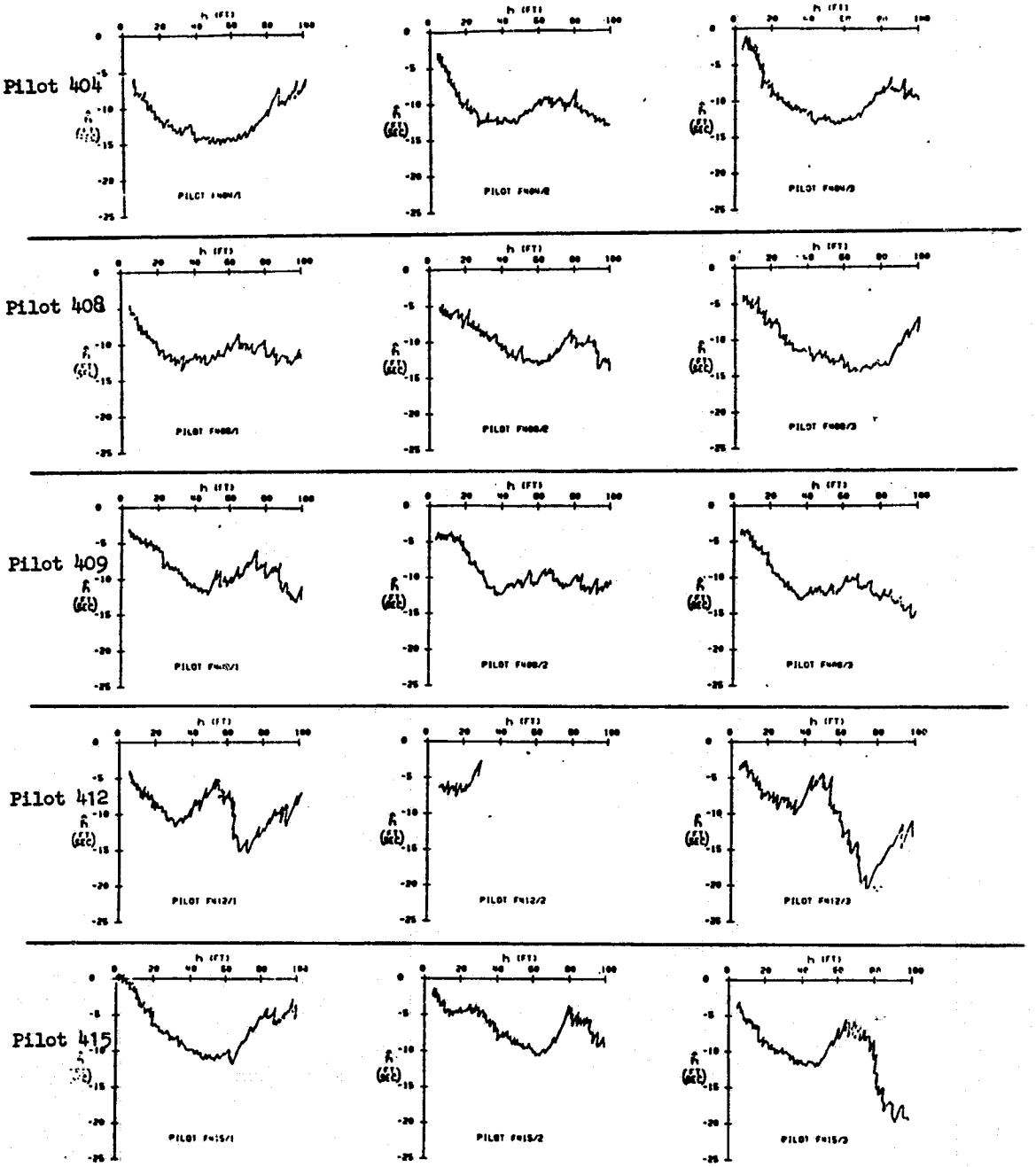


Figure 2. Samples of Landing Trajectory Phase Planes from Actual Landings (three check-ride landings for five of the 13 flight-trained pilots)

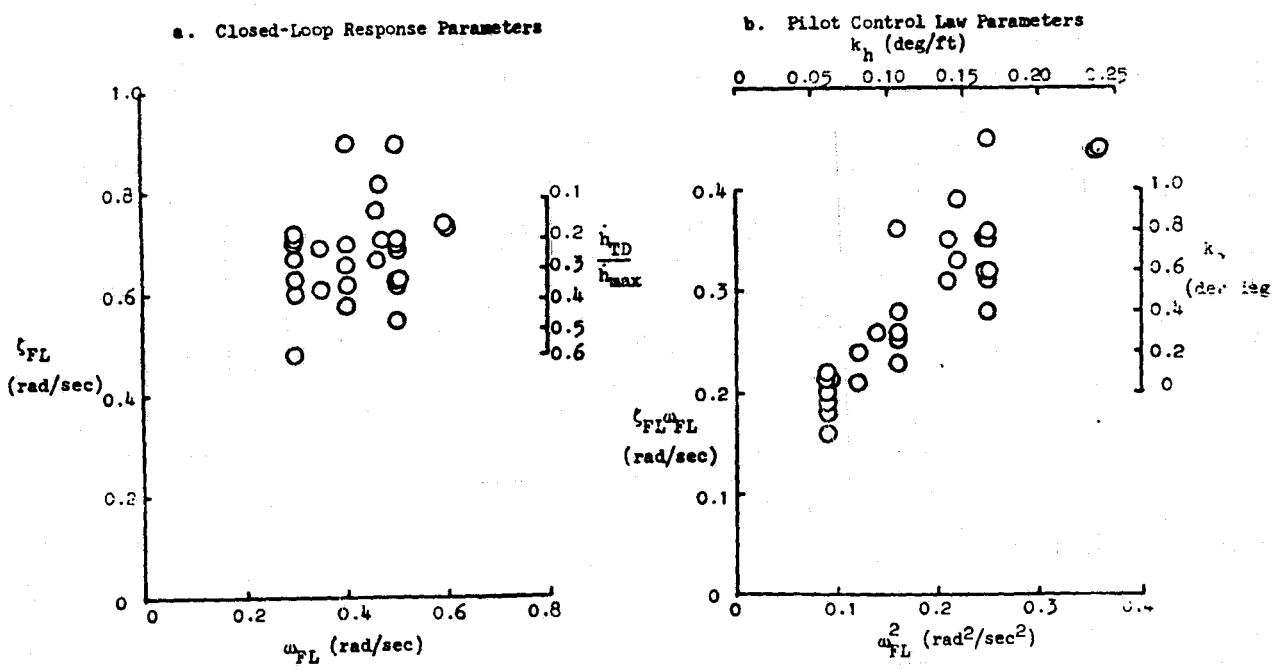


Figure 3. Closed-Loop Flare Parameters Representative of Skilled Pilots in Flight

TABLE 1. AVERAGE PILOT BEHAVIOR IN FLIGHT — PILOTING TECHNIQUE

Features Of Maneuver and Pilot Behavior	Parameters	Flight-Trained Pilots Exhibiting Good Landings*	Remarks
Control of touchdown sink rate	ζ_{FL}	0.68 ± 0.09†	Effective reduction in sink rate
	$\dot{h}_{TD}/\dot{h}_{max}$	0.25 ± 0.14	
Abruptness of flare maneuver	ω_{FL} (rad/sec)	0.42 ± 0.09	Bandwidth high enough to precede airspeed decay (about 0.1 rad/sec) and low enough to accommodate lag in pitch attitude command (about 1 rad/sec)
	ω_{c_h} (rad/sec)	0.28 ± 0.06	
Height Feedback	ω_{FL}^2 (rad^2/sec^2)	0.19 ± 0.08	Consistent with flight manual — about 3.5 deg attitude change over the final 30 to 40 ft
	k_h (deg/ft)	0.13 ± 0.05	
Direction-of-flight feedback	$\dot{\zeta}_{FL}\omega_{FL}$ (rad/sec)	0.29 ± 0.08	Significant feedback of direction-of-flight or its equivalent
	k_γ (deg/deg)	0.45 ± 0.35	

* $|\dot{h}_{TD}| < 5$ ft/sec, no floating (Group FA).

† Mean ± standard deviation.

final reduction in sink rate. In fact the duck-under maneuver fitted the phase plane trajectory of the flare itself, i.e., the same pilot control law generated both maneuver segments as shown in Fig. 4. Furthermore the initiation of the duck-under ranged even greater in height than did the flare, per se.

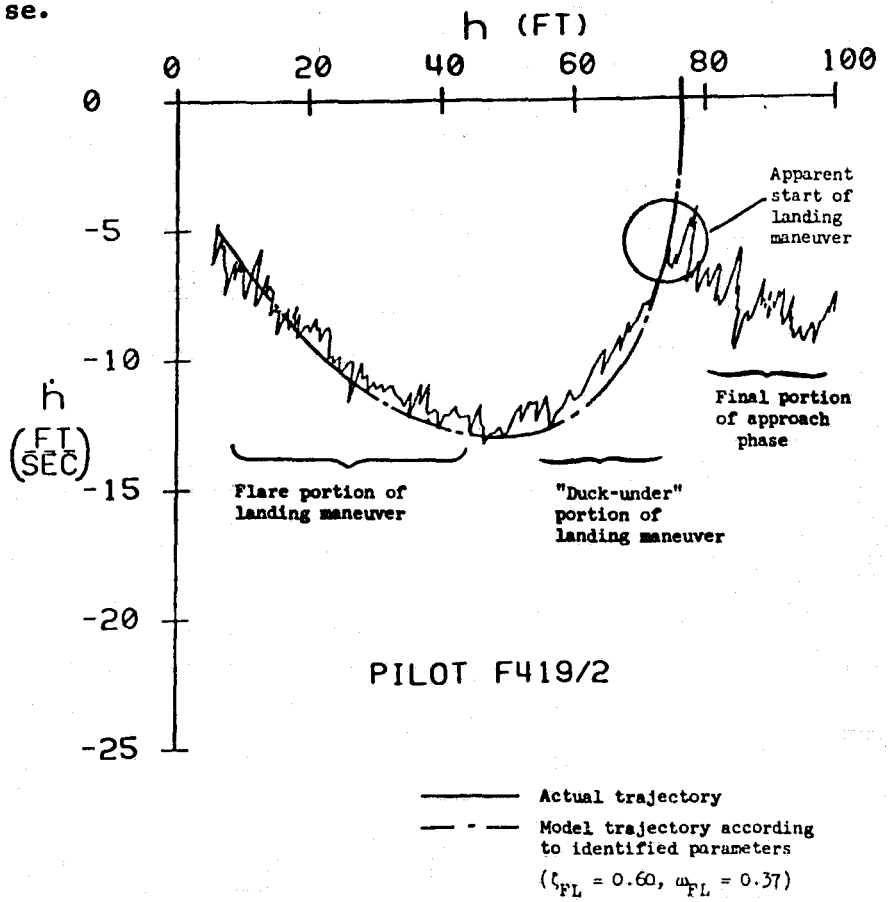


Figure 4. Typical Landing Maneuver Performed in the Actual Aircraft

A NASA research pilot observed that the combined duck-under and flare was a natural action aimed at adjusting touchdown point along the runway. A normal electronic glide slope may intercept the runway at a conservative distance from the runway threshold. Thus at some judicious point prior to flare the pilot may elect to transition from the electronic glide slope to a lower approach slope. The data suggest that this final adjustment is integrated with the flare and that the height of that adjustment corresponds to how large a change is desired in the nominal touchdown point. Hence "flare height" or "landing initiation height" should not be regarded as a constant. Rather it is a "control" used to alter the point of touchdown along the runway.

Landings in the Simulator

Substantially different average landing behavior was observed in the simulator. The same analysis procedure was applied, i.e., identification of landing maneuver parameters from phase plane trajectories. A summary of results is given in Table 2.

TABLE 2. AVERAGE PILOT BEHAVIOR IN FLIGHT AND IN THE SIMULATOR — SIMULATOR FIDELITY

Features of Maneuver and Pilot Behavior	Parameters	All Flight-Trained Pilots in Flight	All Pilots in Flight	All Pilots in the Simulator	Remarks
Control of touchdown sink rate	ζ_{FL}	0.68	0.67	0.58	Harder landings in the simulator
	$\dot{h}_{TD}/\dot{h}_{max}$	0.25	0.27	0.42	
Abruptness of flare maneuver	ω_{FL} (rad/sec)	0.40	0.37	0.36	No difference
	ω_{c_h} (rad/sec)	0.27	0.25	0.27	
Height Feedback	ω_{FL}^2 (rad^2/sec^2)	0.17	0.13	0.15	No difference
	k_h (deg/ft)	0.11	0.08	0.10	
Direction-of-flight or sink rate feedback	$\zeta_{FL}\omega_{FL}$ (rad/sec)	0.28	0.25	0.20	Direction-of-flight loop lacking in the simulator
	k_γ (deg/deg)	0.40	0.25	0	

The most obvious difference between simulator and flight was in the firmness of landings. This is reflected in the ratio of sink rate decay, h_{TD}/h_{max}, and the effective damping ratio, ζ_{FL}. At the same time the abruptness of the flare maneuver and corresponding height feedback were comparable between simulator and flight.

One important piloting technique implication from the above observations is that the effective direction-of-flight or sink rate feedback is inadequate in simulator landings. A further implication is that a cue deficiency exists. The exact nature of that cue was, however, not clear from the data although visual perception of sink rate is suspected.

An additional feature of the simulator data was that there was an absence of the initial duck-under maneuver which was so prominent in the flight data. This could be interpreted as either the absence of a runway distance cue or a different approach geometry which made a duck-under unnecessary.

As suggested in Ref. 4 the fidelity of the simulator can be judged on the net difference in piloting technique exhibited in the simulator compared to actual flight. Therefore, for the case reported here, one could question the fidelity of the essential sink-rate or direction-of-flight cues as well as distance along the runway. At the same time it is not clear, without further experimentation, what the specific origins of these difficulties were in engineering terms used to specify simulator components.

CONCLUDING REMARKS

The airline landing data analyzed yielded a rich variety of results with implications in several areas including quantification of piloting technique and the fidelity of an airline training simulator for the landing maneuver. Besides providing important quantification in these various areas, the data have also provided the basis for a revised analytical model of the flare maneuver. In fact the model developed provides a useful bridge between the raw data collected and the ensuing interpretations of those data.

Several metrics have evolved with regard to describing the landing maneuver. The first metric is the phase plane representation to characterize the flare maneuver, not only in terms of the ultimate landing performance but also how that performance was achieved: whether it was the result of a last-minute abrupt pull-up leaving no room for error or misjudgment, or whether it was the result of an exceedingly gentle decay in sink rate which might be accompanied by a large loss of airspeed prior to touchdown. The phase plane also, of course, shows where there were dangerously high sink rates at low altitudes or if there was a floating or ballooning tendency.

Two metrics which bear a direct dependence upon the effective closed-loop parameters are the inferred pilot-vehicle loop gains, namely the height gain and the direction-of-flight gain. The height gain was shown to be dependent upon the closed-loop natural frequency and the true airspeed. The direction-of-flight gain was shown to be a function of the product of damping ratio and natural frequency along with the basic vertical response lag for the aircraft.

An important aspect of the analysis performed here is the quantification of the landing maneuver as it is performed on the actual aircraft. This provides an important baseline for examining simulator fidelity. Without this description of piloting technique, we would have to rely far more heavily upon terminal landing performance (i.e., scoring of the touchdown sink rate or distance along the runway) or on strictly subjective judgments.

There are indications in the closed-loop pilot-vehicle response parameters that the fidelity of the training simulator used in this study was deficient in at least one modality. The outside visual scene is most suspect; but the simulator motion system cannot be ruled out without further investigation nor, for that matter, can the simulator model implementation.

REFERENCES

1. Heffley, Robert K., Schulman, Ted M., Clement, Warren F., and Randle, Robert J., Jr., An Analysis of Airline Landing Data Based on Flight and Training Simulator Measurements, Systems Technology, Inc., Technical Report No. 1172-1 (forthcoming NASA CR), June 1981.

2. Randle, Robert J., Jr., Tanner, Trieve A., Hamerman, Joy A., and Showalter, Thomas H., The Use of Total Simulator Training in Transitioning Air-Carrier Pilots: A Field Evaluation, NASA TM 81250, January 1981.

3. McRuer, D. T., and Krendel, E. S., Mathematical Models of Human Pilot Behavior, AGARD-AG-188, January 1974.

4. Heffley, Robert K., Clement, Warren F., Ringland, Robert F., et al., Determination of Motion and Visual System Characteristics for Flight Simulation, Systems Technology, Inc., Technical Report No. 1162-1, April 1981.

5. Key, David L., (Ed.), Fidelity of Simulation for Pilot Training, AGARD Advisory Report No. 159, October 1980.

6. Smith, Jon M., Mathematical Modeling and Digital Simulation for Engineers and Scientists, John Wiley and Sons, New York, 1977, pp. 195-199.

7. White, Maurice D., Proposed Analytical Model for the Final Stages of Landing a Transport Airplane, NASA TN D-4438, April 1968.

8. Hoh, Roger H., Craig, Samuel J., and Ashkenas, Irving L., Identification of Minimum Acceptable Characteristics for Manual STOL Flight Path Control. Volume III: Detailed Analyses and Tested Vehicle Characteristics, FAA-RD-75-123, III, June 1976.

9. Bray, Richard S., A Piloted Simulator Study of Longitudinal Handling Qualities of Supersonic Transports in the Landing Maneuver, NASA TN D-2251, April 1964.

10. McRuer, Duane, Ashkenas, Irving, and Graham, Dunstan, Aircraft Dynamics and Automatic Control, Princeton University Press, Princeton, New Jersey, 1973.

omit

Model Analysis of Delay Compensation Schemes for Simulators

Sheldon Baron

Bolt Beranek and Newman Inc.
10 Moulton Street
Cambridge, MA 02238

Unwanted delays in digitally-based flight simulators are inherent and arise from the basic sample-data nature of the simulation, the algorithms employed and the lags generated in simulating the appropriate cue environment. With very high speed powerful computers, the latter source of delay tends to dominate, particularly with high resolution computer generated imagery being employed. It can be shown theoretically and has been demonstrated experimentally in manual control tasks, that transport delays in a feedback control situation will degrade performance. Moreover, it is also recognized that in these situations, delays can increase pilot workload and alter pilot control strategy (so as to avoid PIO problems). For these reasons, it is important to minimize simulation hardware and software delays in the design process and, subsequently, to compensate for any remaining significant delay, if possible.

Of course, the significance of a particular simulation delay and the improvement provided by a given compensator will depend on the particulars of the simulated flight task, i.e., on the vehicle dynamics, disturbances and control objectives as well as on the ability of the pilot to compensate for the delay. This complexity, and the importance of considering pilot response in evaluating the effects of delays and delay compensation schemes, suggests that a closed-loop pilot-vehicle model would be a useful tool in analyzing such problems.

This paper describes an on-going effort to apply the Optimal Control Model (OCM) to the problems of design and evaluation of delay compensation filters. A simple roll tracking task is considered and rms tracking performance as a function of simulator delay is computed. System crossover frequency and pilot phase angle near crossover are also computed. The OCM results compare favorably with those obtained in an independent experimental investigation.

The OCM is then used to synthesize a compensation filter by developing a rational approximation to the optimal predictor that is incorporated in the OCM. This compensator is evaluated using the OCM and the results predicted by the model suggest a restoration of performance to undelayed levels. However, it should be noted that the compensator has not been tested experimentally at this time.

omit

PRECEDING PAGE BLANK NOT FILMED

A MANNED SIMULATOR INVESTIGATION OF THE EFFECTS OF AN INTEGRATED ISOMETRIC CONTROLLER ON PILOT WORKLOAD FOR HELICOPTER NAP-OF-THE-EARTH FLIGHT

Edwin W. Aiken
and
Christopher L. Blanken

Aeromechanics Laboratory
U.S. Army Research and Technology Laboratories (AVRADCOM)
Ames Research Center
Moffett Field, California

John C. Hemingway

NASA-Ames Research Center
Moffett Field, California

ABSTRACT

The design and interim results of an experiment conducted to assess the effects on pilot workload of replacing one or more of the standard helicopter controllers with an integrated isometric side-stick controller are described. A fixed-base piloted simulation of low-altitude helicopter maneuvers typical of daytime nap-of-the-earth flight is used to investigate the effects of certain critical parameters, including controller orientation, controlled vehicle dynamic characteristics, and the number of axes controlled through the isometric device. In addition to subjective measures of these effects through Cooper-Harper pilot ratings and pilot commentary, objective measures of pilot workload are collected. Control activity provides a measure of physical workload; mental workload is assessed using a secondary task based upon the Sternberg item recognition technique. This task, presented visually through the pilot's head-up display, is implemented in a fashion similar to that used in an in-flight investigation of control/display effects on pilot workload for fixed-wing aircraft landing approaches (reference 1). Results to date indicate the feasibility of using the isometric device for all four controlled axes if appropriate stability and control augmentation is also provided.

REFERENCES

1. Schiflett, S.G.: Evaluation of a Pilot Workload Assessment Device to Test Alternate Display Formats and Control Handling Qualities. Naval Air Test Center Report SY-33R-80. 1980.

IDENTIFICATION OF MULTILOOP PILOT DESCRIBING FUNCTIONS OBTAINED FROM SIMULATED APPROACHES TO AN AIRCRAFT CARRIER

Wayne F. Jewell

Systems Technology, Inc.

SUMMARY

This paper presents some predicted results of a simulation of the carrier aircraft pilot's approach control strategy in the presence of pilot remnant. The aircraft dynamics and the turbulence environment are representative of a trainer-type aircraft. The Non-Intrusive Pilot Identification Program (NIPIP) was used to identify the pilot's control strategy required by this highly-coupled, multiloop control task. The results are presented in terms of frequency responses of the individual elements of the pilot's control strategy and indicate that NIPIP can identify the pilot's describing functions even in the presence of significant amounts of pilot remnant. The next step is to apply NIPIP to a real time, piloted simulation of the same control task. This is planned for the Visual Technology Research Simulator at the Naval Training Equipment Center in Orlando, Florida.

INTRODUCTION

The Non-Intrusive Pilot Identification Program (NIPIP) was developed, evaluated, and applied to a simulated approach and landing of a conventional takeoff and landing (CTOL) aircraft (Refs. 1 and 2). The performance evaluation of NIPIP demonstrated that accurate, unbiased estimates of the pilot's input-output describing function, Y_p, could be obtained without explicit knowledge of the disturbance function (Refs. 1 and 2). The atmospheric disturbance was injected into the aircraft dynamics (instead of being injected into the control loop) and was not used as an input to NIPIP. NIPIP obtained the pilot's input-output describing function by using a time-domain model of Y_p and a least-squares identification algorithm. Furthermore NIPIP used a "sliding" time window to estimate Y_p enabling it to identify time-varying behavior in the pilot's control strategy.

Although the CTOL control task of Refs. 1 and 2 was a multiloop manual control problem, there was very little coupling between the fast/slow (throttle) axis and the flight director (column) axis. In effect, two single axis manual control tasks were being performed simultaneously. In the present case of carrier landing, however, the block diagram in Fig. 1 depicts the highly-coupled, multiloop manual control task used by Navy pilots for final approach. The prescribed Navy piloting technique for controlling the aircraft is to regulate glide slope deviation, d, with throttle, δ_T, and angle of attack, α, with commanded pitch attitude, θ_c. Commanded pitch attitude, in turn, is regulated with the elevator, δ_E, as shown in Fig. 1 (Ref. 3). In reality, however, a pilot learns that he must "crossfeed" the controls in order to "stay ahead" of the aircraft; that is, when the pilot makes a correction to d using δ_T he also adjusts θ_c, as shown in Fig. 1.

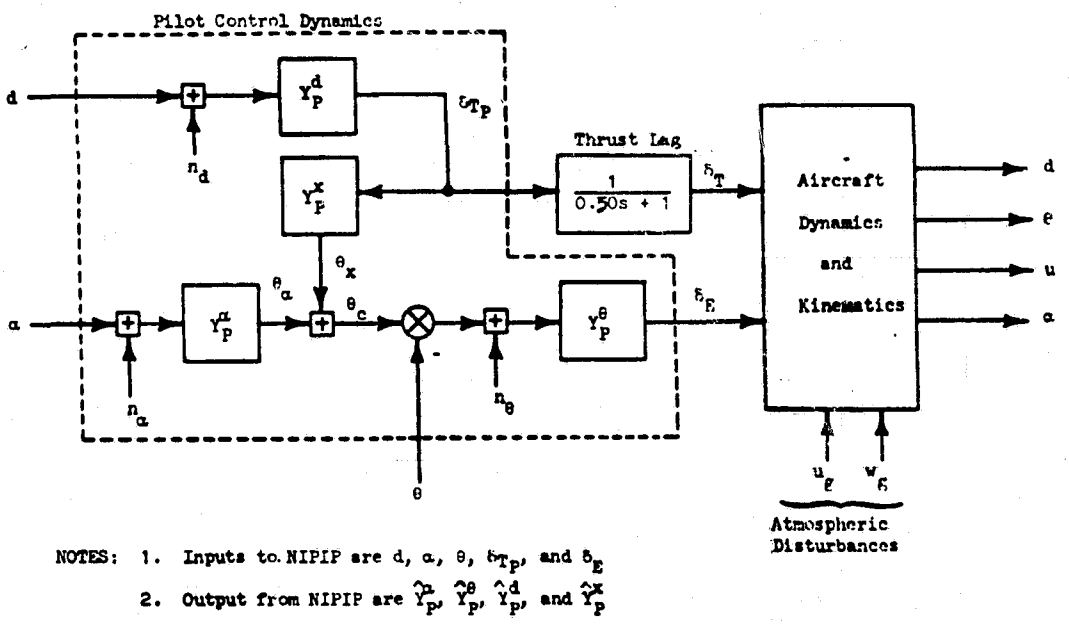


Figure 1. Manual Control Task for STOL Approach and Landing

The analytical study by Heffley, et al., in Ref. 4 demonstrates that the "pursuit-crossfeed" piloting technique described above, and shown in Fig. 1, must be used in order to obtain adequate approach precision. However significant skill development is required to adopt the proper crossfeed for the aircraft and approach speed being flown. If the pilot is required to fly a different aircraft, as is done when proceeding from a trainer to a fleet aircraft, he must readapt his crossfeed gain. Furthermore Ref. 4 shows that compensatory Y_P^α can become a very low and almost negligible gain when the pilot has learned to develop the pursuit crossfeed, Y_P^x.

To prepare for identifying skill development in the pilot's control strategy, NIPIP was used first in an inanimate simulation in scaled time to quantify the various elements of the pilot's control technique; i.e., Y_P^d, Y_P^x, Y_P^α, and Y_P^θ in Fig. 1 in the presence of pilot remnant. Our ultimate goal is to use NIPIP for analyzing simulated approaches to an aircraft carrier in real time using the Visual Technology Research Simulator (VTRS, Ref. 5) at the Naval Training Equipment Center (NTEC) in Orlando, Florida. This will be done within the near future using Navy pilots and a number of simulated aircraft.

Prior to using NIPIP in conjunction with the piloted simulation described above, we wanted to know if NIPIP could indeed identify the individual elements of the complex control loop structure shown in Fig. 1. In order to prove NIPIP's ability to do this, we simulated the aircraft dynamics, the pilot describing functions, and the pilot remnant shown in Fig. 1. The combined pilot-aircraft system was disturbed with simulated atmospheric turbulence. The results of analyzing this data are presented in the remainder of this paper.

TECHNICAL APPROACH

The aircraft dynamics used in the simulation were representative of a T2-C at 108 kt, flaps fully extended (Ref. 6). The various pilot describing functions indicated in Fig. 1 were as follows:

$$Y_p^{\theta} = -0.30 \frac{s + 10.0}{s + 3.0} \left(\text{rad } \delta_E/\text{rad } \theta_e\right)$$

$$Y_p^{d} = 32.6 \frac{2.0}{s + 2.0} \left(\text{lb } \delta_T/\text{ft d}\right)$$

$$Y_p^{x} = 8.2\text{E-}5 \frac{3.0}{s + 3.0} \left(\text{rad } \theta_x/\text{lb } \delta_{T_P}\right)$$

$$Y_p^{\alpha} = 0.0$$

The compensation defined above yields a pitch-loop bandwidth of about 3.0 rad/sec and a glide slope loop bandwidth of about 0.50 rad/sec. The angle of attack loop was not closed for the results presented herein (it is a very low bandwidth loop) but will be added in the future. The crossfeed was designed such that the steady-state airspeed will remain unchanged for any amount of throttle deflection. It should be mentioned that, from a control system design point of view, it is possible to use pure gains for Y_p^{θ}, Y_p^{d}, and Y_p^{x}. However this would have been a trivial identification problem for NIPIP and would not necessarily be representative of human pilot dynamics.

Pilot remnant was simulated using shaped white noise and injected into the control loops as shown in Fig. 1. The shaping filters used to obtain n_{θ} and n_d were (Ref. 7)

$$\frac{n_d}{\eta_d} = \frac{2.58}{s + 3.33} \sigma_{n_d}$$

$$\frac{n_{\theta}}{\eta_{\theta}} = \frac{1.41}{s + 1.0} \sigma_{n_{\theta}}$$

where η_d and η_{θ} are independent white noise sources and σ_{n_d} and $\sigma_{n_{\theta}}$ are the rms values of n_d and n_{θ}. The values of σ_{n_d} and $\sigma_{n_{\theta}}$ were set such that

$$\sigma_{n_d} \doteq 0.25\, \sigma_d$$

$$\sigma_{n_{\theta}} \doteq 0.25\, \sigma_{\theta}$$

where σ_d and σ_{θ} are the rms values of d and θ when no remnant is present.

The atmospheric turbulence was simulated using pseudo-random noise (sum of sine waves) to simulate the axial, u_g, and the vertical, w_g, gusts. The

rms values for both u_g and w_g were 3.0 fps for the results contained herein. The bandwidths of both u_g and w_g were about 0.50 rad/sec.

RESULTS

Some typical time histories of the pilot-aircraft response to turbulence with and without simulated pilot remnant are shown in Fig. 2. Note that the remnant has a fairly large effect on the controls, δ_T and δ_E, but a relatively small effect on the aircraft response. This is because the aircraft acts as a filter and smooths the noisy control inputs. The δ_T response looks granular because a sampling rate of 0.50 sec was deliberately used to simulate Y_p^d. A sampling rate of 0.10 sec was used to simulate Y_p^θ.

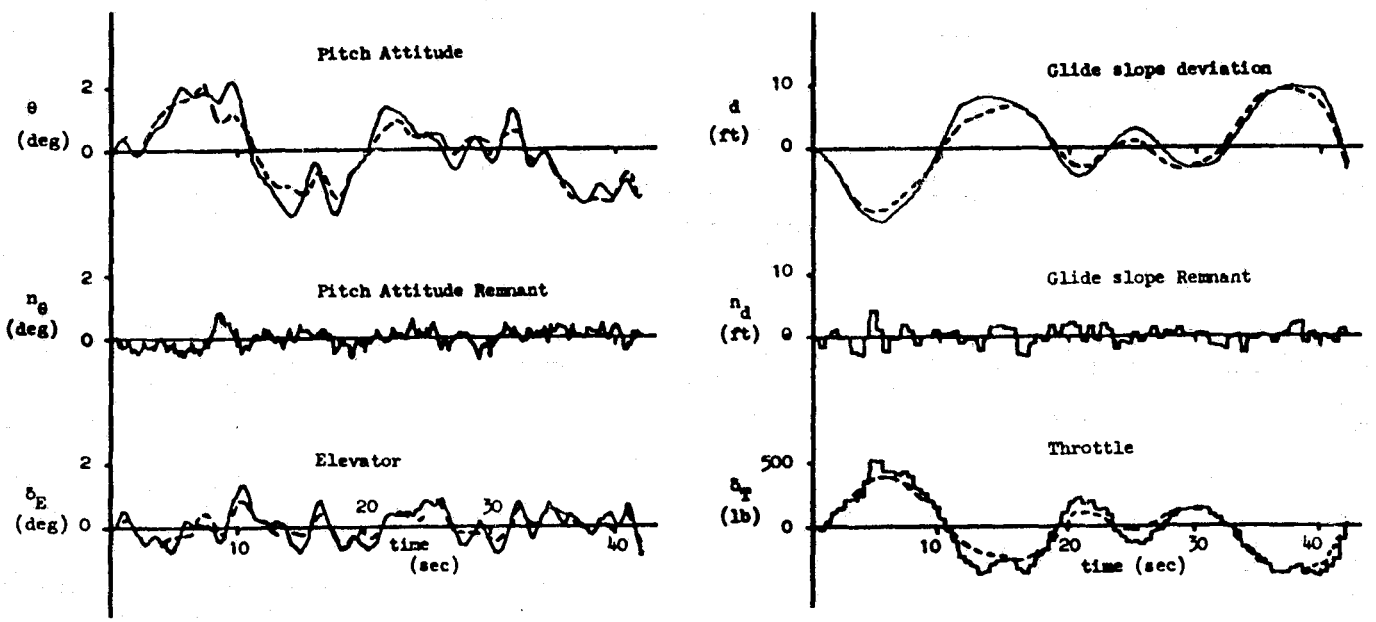


Figure 2. Pilot-Aircraft Response to Turbulence with and Without Simulated Pilot Remnant

The time histories of d, δ_T, θ, and δ_E shown in Fig. 2 were used as inputs to NIPIP, from which NIPIP computed the desired pilot describing functions, Y_p^θ, Y_p^x, and Y_p^d. A detailed discussion of how NIPIP performs these calculations can be found in Ref. 2 and will not be repeated here.

Figures 3, 4, and 5 compare the frequency responses of the actual pilot dynamics to the outputs of NIPIP. The length of the time windows used in computing Y_p^θ and Y_p^x was 30 sec, but was 60 sec for Y_p^d. The longer time window was used for computing Y_p^d because the $d \rightarrow \delta_T$ loop has a much lower bandwidth than the $\theta \rightarrow \delta_E$ loop. The three estimates of $Y_p(j\omega)$ shown in the figures demonstrate how the NIPIP outputs vary with time (even though the simulated pilot describing functions were stationary). The variation of all three estimates is fairly low, especially in the neighborhood of the crossover frequency. The explanation for these phenomena is the more adverse signal-to-noise ratios outside the region of the crossover frequency. This result was also demonstrated in Ref. 2.

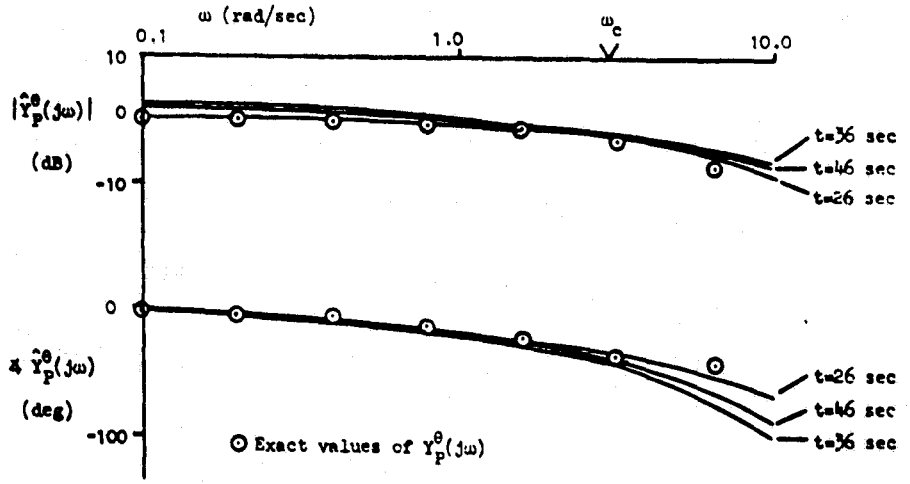


Figure 3. Frequency Response of $\hat{Y}_\theta$ with 25 Percent Remnant and a Time Window of 30 sec

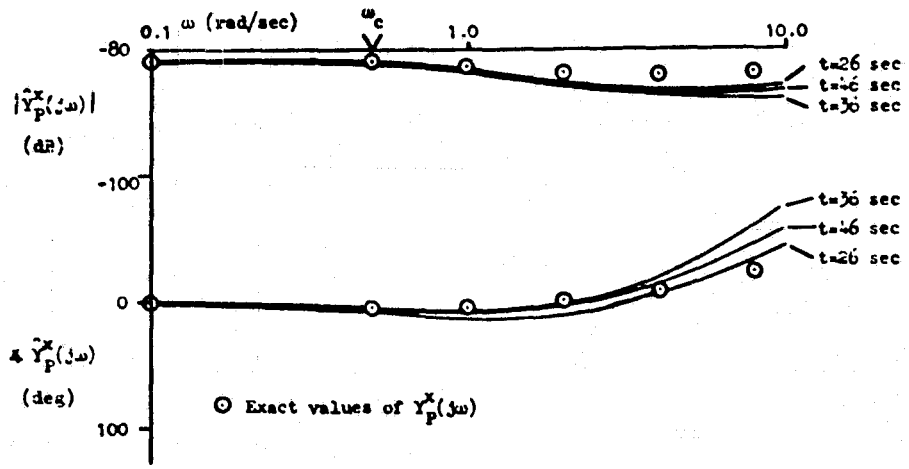


Figure 4. Frequency Response of $\hat{Y}_p^x$ with 25 Percent Remnant and a Time Window of 30 sec

CONCLUDING REMARKS

The Non-Intrusive Pilot Identification Program (NIPIP) was used to estimate the pilot's control strategy required for the final approach and landing on an aircraft carrier. The estimates for the pilot's describing functions are quite accurate in the region of their respective crossover frequencies, (i.e., ω_d and ω_θ). The errors could be further reduced by increasing the lengths of the time window. The penalty for doing this, however, is that any time-variation in the pilot's actual control strategy tends to be masked. The issue of time-variation in piloting technique will be addressed when we analyze the data from a piloted simulation on the Visual Technology Research Simulator at the Naval Training Equipment Center in Orlando, Florida.

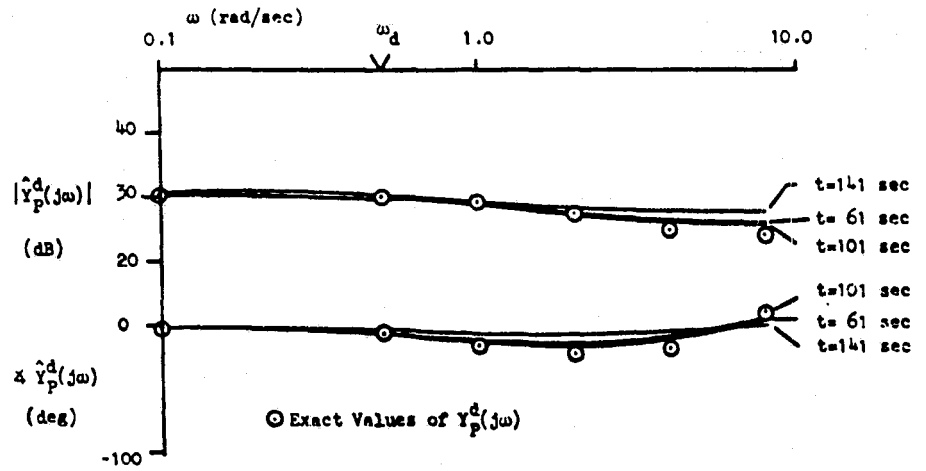


Figure 5. Frequency Response of $\hat{Y}_p^d(j\omega)$ with 25 Percent Remnant

REFERENCES

1. Jewell, Wayne F., "Application of a Pilot Control Strategy Identification Technique to a Joint FAA/NASA Ground-Based Simulation of Head-Up Displays for CTOL Aircraft," Sixteenth Annual Conference on Manual Control, MIT, Cambridge, Mass., May 5-7, 1980, pp. 395-409.

2. Jewell, Wayne F., and Schulman, Ted M., A Pilot Control Strategy Identification Technique for Use in Multiloop Control Tasks, Systems Technology, Inc., Technical Report No. 1153-2 (forthcoming NASA Contractor Report), May 1980.

3. Anon., Undergraduate Pilot Training Task Analysis, Phase II Report, Hq. Chief of Naval Air Training, NAS, Corpus Christi, Texas, July 1975.

4. Heffley, Robert K., Clement, Warren F., and Craig, Samuel J., "Training Aircraft Design Considerations Based on the Successive Organization of Perception in Manual Control," Sixteenth Annual Conference on Manual Control, MIT, Cambridge, Mass., May 5-7, 1980, pp. 119-127.

5. Chambers, Walt S., "AWAVS: An Engineering Simulator for Design of Visual Flight Training Simulators," J. of Aircraft, Vol. 14, No. 11, November 1977, pp. 1060-1063.

6. Buenz, D. A., et al., Identification of T-2 Aerodynamic Derivatives from Flight Data, Systems Control Inc., for Naval Air Development Center, March 1975 (AD-A021 996).

7. McRuer, D. T., and Krendel, E. S., Mathematical Models of Human Pilot Behavior, AGARD-AG-188, January 1974.

FURTHER TESTS OF A MODEL-BASED SCHEME FOR PREDICTING PILOT OPINION RATINGS FOR LARGE COMMERCIAL TRANSPORTS

by

W. W. Rickard
McDonnell Douglas Corporation
Douglas Aircraft Company
Long Beach, California

and

W. H. Levison
Bolt Beranek and Newman Inc.
Cambridge, Massachusetts

ABSTRACT

A methodology is demonstrated for assessing longitudinal-axis handling qualities of transport aircraft on the basis of closed-loop criteria. Six longitudinal-axis approach configurations were studied covering a range of handling quality problems that included the presence of flexible aircraft modes. Using closed-loop performance requirements derived from task analyses and pilot interviews, predictions of performance/workload tradeoffs were obtained using an analytical pilot/vehicle model. A subsequent manned simulation study yielded objective performance measures and Cooper-Harper pilot ratings that were largely consistent with each other and with analytic predictions.

INTRODUCTION

A certain dichotomy is associated with the assessment of flying qualities. From the pilot's point of view, the flying qualities of an airplane, in a given task, relate to the degree to which satisfactory performance can be achieved with reasonable workload levels. Nevertheless, flying quality specifications are written in terms of open-loop vehicle response characteristics to help the airplane manufacturer comply with the specifications. Accordingly, considerable effort has been expended to find the combination of aircraft response parameters that will reliably predict task performance and pilot workload.

In contrast to open-loop vehicle-centered criteria, pilot/vehicle model analysis allows one to explore issues related to closed-loop performance as well as to workload demands made on the pilot. The effects of external disturbances and control/display parameters, as well as inherent pilot limitations, can be considered. Perhaps most important, predictive schemes based on pilot/vehicle analysis are not constrained to "conventional" dynamics and can therefore be applied to flying quality studies of aircraft having high-order reponse characteristics.

Hess[1] and Levison[2, 3] have proposed similar schemes, based on the optimal control model (OCM) for pilot/vehicle systems, for predicting pilot-opinion ratings. Levison's scheme was recently tested against data obtained in a previous simulation study of commercial transport handling qualities.[4] Results of this test were sufficiently encouraging to warrant further exploration of the methodology.

This paper summarizes the results of a subsequent study by Bolt Beranek and Newman Inc. and Douglas Aircraft Company in which the analytic scheme for assessing longitudinal axis handling qualities of commercial transport aircraft in the landing approach was rigorously tested. Study goals included development of closed-loop performance criteria, a tightly constrained manned simulation to yield Cooper-Harper opinion ratings[5] with minimal interpilot variability, and compilation of a data base of objective performance measures suitable for methodological development. An additional goal was to explore the effects of simulating flexible modes of transport aircraft, and to determine whether or not the analytic scheme would predict these effects.

The prediction scheme is based on the following assumptions:

1. Pilot rating is a function of the flight task
2. For a given flight task, one or more critical subtasks exist which serve as the primary determinants of pilot ratings
3. Performance requirements are well defined for each critical subtask
4. Pilot opinion is based partly on the degree to which desired performance is achieved and partly on the information-processing workload associated with the task
5. A reliable model exists for predicting performance/workload tradeoffs for relevant flight tasks.

These assumptions lead to the procedure diagrammed in Figure 1. In effect, the analytic prediction scheme parallels the procedure that would be followed in performing a well-controlled handling quality simulation study, the major difference being use of the optimal control pilot/vehicle model

rather than a human to obtain pilot ratings and other performance measures.

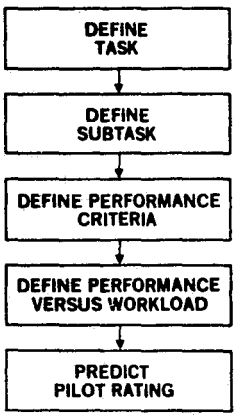


FIGURE 1. PROCEDURE FOR PREDICTING PILOT RATING

The following rating expression was developed in the preceding study and used in the current effort:

$$R = 10\left[\frac{\sigma}{\sigma+\sigma_o}+\frac{A}{A+A_o}\right] \qquad (1)$$

$$1 \leq R \leq 10$$

where R is the predicted Cooper-Harper pilot rating, σ is the probability that one or more important system variables will exceed its maximum acceptable value, and A is a measure of the relative attention (i.e., workload) associated with the task. The pilot is assumed to operate on the performance/workload tradeoff curve, predicted by the OCM, so as to minimize R.

A good fit with the experimental data of the preceding study was found using $\sigma_o = 0.1$ and $A_o = 2$. (Since "attention" is considered relative to that appropriate to a standardized laboratory tracking task[6, 7] rather than to some assumed capacity, values greater than unity are possible.)

The paper is organized into the following topic areas: (1) task definition, (2) pre-experiment analysis to select experimental configurations and predict aircraft handling qualities, (3) description of the experiment, (4) experimental results and analysis, and (5) conclusions.

TASK DEFINITION

The flight tasks to be performed by the test pilots were defined and closed-loop criteria were specified for model analysis. As in the preceding study,[2] the task was that of piloting a simulated large commercial transport aircraft in the final approach. Three subtasks were defined: (1) altitude stationkeeping prior to glideslope capture, (2) glideslope capture, and (3) post-capture tracking of the glideslope. Flare and landing were not considered, and were not performed in the simulation study. For purposes of pre-experimental model analysis, zero-mean turbulence as defined by the Dryden model[8] was assumed, with q and r components omitted.

To ascertain closed-loop requirements, interviews were held with five potential test pilots to determine what they considered to be maximum acceptable values, or "limits," for important system variables in moderate turbulence. (In general, the pilots interpreted a "maximum" value as an excursion indicative of poor approach performance.) Assessments were obtained for each of the flight subtasks, and for various altitudes with respect to the glideslope tracking subtask. Since model analysis was performed for frozen-point conditions appropriate to glideslope tracking at an altitude of 500 feet, the subjective acceptable excursions for that condition are reported here.

On the average, the following zero-to-peak allowable excursions from trim were specified:

Glideslope: 1/2 dot
Sink rate: 250 feet per minute
Airspeed: 7.6 knots
Pitch: 3.5 degrees
Stick: 28 percent of maximum excursion
Thrust: 4 percent of aircraft weight

For airspeed and sink rate excursions, which had asymmetric criteria, the above values reflect one-half the distance between upper and lower bounds. The limit on thrust represents a distillation of the pilot responses, which were expressed in different units (inches throttle movement, percent N_1, change in EPR) by different pilots. The pilots agreed that there was also a subjective limit to pitch rate but, as they could not assign a quantitative value to this parameter, it was excluded from the list of performance requirements.

Although the pilots did not provide subjective limitations to rate-of-change of stick and throttle, "limits" for these quantities were defined partly to satisfy certain mathematical requirements of the optimal control model and partly to satisfy physical constraints. A stick rate limit of 28 percent maximum slew rate was assumed, and the limit on rate-of-change of thrust was set equal to one-half the limit on thrust deviation to reflect the low bandwidth operation of this control.

To provide the scalar quadratic performance index needed to obtain model solutions, weighting coefficients were defined as the reciprocals of the squares of maximum allowable values. Thus, an rms deviation of a given system variable equal to its "limit" contributed one unit to the overall "cost."

PRE-EXPERIMENT ANALYSIS

This phase of the study consisted of two tasks: (1) preliminary selection of candidate aircraft configurations and (2) pre-experimental model analysis. The objectives of the latter task were to select six configurations for the simulation study and to obtain predictions of pilot ratings and closed-loop performance measures to allow rigorous testing of the analytic methodology.

PRELIMINARY SELECTION OF CANDIDATE CONFIGURATIONS

The pilot rating data from the test reported by Rickard[4] were used to demonstrate the feasibility of using the OCM to estimate pilot ratings. A thorough validation of this procedure, however, requires comparison not only of the pilot rating data but also of the workload and tracking performance data for human pilots with model predictions. These data had not been recorded in that test. Thus, another test was planned in which all the needed data would be recorded.

The first task undertaken in designing the test was the selection of a set of configurations to be evaluated. A primary goal was to produce data with high statistical reliability. This meant that many replications of the test matrix, or repeated evaluations of each configuration, would be needed. This meant that the test matrix would have to be small to keep costs reasonable.

One would prefer a large test matrix so one could evaluate the model against a wide range of airplane characteristics. Since the test matrix had to be kept small, it was decided that the configurations should be chosen to vary the most important properties. Among the issues considered critical in flying qualities today are relaxed static stability, control augmentation, and structural dynamics. A set of eight configurations was designed which varied these properties.

The flying qualities of the eight configurations, as predicted by existing criteria, are shown in Table 1. There are columns for five -8785 criteria: (1) short period frequency versus acceleration sensitivity ($\omega_{n_{sp}}$ versus n/α), (2) short period damping (ζ_{sp}), (3) phugoid damping (ζ_{ph} or $T_{2_{ph}}$), (4) static stability, and (5) flight path stability ($d\gamma/dV$). The pilot, of course, cannot be asked to rate these individual criteria; his rating of longitudinal flying qualities represents their sum. Since the -8785 provides no guidance on how to combine the pieces, one must use his own judgment. The judgment used here was to represent the "-8785 OVERALL" as the worst of the five preceding columns. The next column, labeled "BANDWIDTH," is a flying quality prediction using a frequency domain pilot-in-the-loop criterion.[4] This criterion has been demonstrated reliably to predict pilot opinion of longitudinal maneuvering dynamics. As such, it is not sensitive to $d\gamma/dV$, which is a measure of long-term flight path response. It was shown by Rickard[4] that the combination of the Bandwidth and flight path stability estimates, labeled BW + $d\gamma/dV$, yields an estimate of longitudinal flying qualities more accurate than the one labeled -8785 OVERALL. The criteria in Table 1 are the tools used to design a matrix of eight configurations for this test. Only six were simulated. Model analysis was used as described in the next section to eliminate two from the test matrix.

Configuration 1 is by time-honored tradition the baseline. According to the estimates, it has Level 1 longitudinal flying qualities. Configurations 8 and 3 explore a progression of increasing static instability, having times to double of 7.7 and 2.4 seconds, respectively. Configuration 2 was chosen to explore the issue of flight path stability, with $d\gamma/dV = 0.34$, where 0.24 is the Level 3 limit. Configurations 4 and 5 explore the issue of control augmentation. They are the same airplane, an advanced supersonic transport, without and with a full-state feedback flight control system which was designed using implicit model following. The unaugmented airplane has very poor flying qualities, while the augmented version has fair to good flying qualities, depending on the criteria used.

TABLE 1
FLYING QUALITY LEVELS OF TEST CONFIGURATIONS 1 THROUGH 8

CONFIG NO.	$\omega_{n_{sp}}$ vs n/α	ζ_{sp}	ζ_{ph} or $T_{2_{ph}}$	STATIC STABILITY	$d\gamma/dV$	-8785 OVERALL	BAND-WIDTH	BW + $d\gamma/dV$
1	1	1	1	STABLE	1	1	1	1
2	1-1/2	1	1	STABLE	WORSE THAN 3	WORSE THAN 3	1	2-1/2
3	WORSE THAN 3	2	WORSE THAN 3	UNSTABLE	1	WORSE THAN 3	3	3
4	WORSE THAN 3	1	WORSE THAN 3	UNSTABLE	3	WORSE THAN 3	3	3
5	2	1	1	STABLE	1	2	1	1
6	3-1/2	2	1	STABLE	1	3-1/2	1	1
7	3	1	1	STABLE	1	3	1	1
8	WORSE THAN 3	1	WORSE THAN 3	UNSTABLE	1	WORSE THAN 3	2	2

The last two configurations, 6 and 7, explore the effect of structural dynamics on flying qualities. Both have the same rigid-body equations, with Configuration 7 having two additional degrees of freedom representing structural dynamics. The criteria indicate that these configurations will be rated the same. The -8785 criteria, which cannot estimate the effect of structural modes, predict Level 3 flying qualities. The short period frequencies are too low and the damping ratio unacceptable or too high. The Bandwidth criterion, which should be able to predict this effect as it makes no assumption about model order, predicts Level 1 flying qualities.

MODEL ANALYSIS

Before obtaining model predictions, it was necessary to specify various independent model parameters relating to the pilot's information-processing limitations. The reader is referred to documentation from the preceding study[2] for methodological details.

Parameters reflecting limitations on the pilot's information-processing capabilities were defined. An effective perceptual threshold was associated with each perceptual variable assumed to be used by the pilot. On the basis of laboratory tracking experiments,[9] thresholds of 0.05 degree and 0.2 degree per second visual arc were assumed, respectively, for perception of the displacement and velocity of a given display indicator. Analysis of the cockpit displays yielded the following perceptual thresholds, in problem units, for an altitude of 500 feet: (a) 4.7 feet height error, (b) 19 feet per second sink rate error, (c) 4.3-degree pitch error, (d) 1.7 degrees per second pitch rate, and (e) 1.9 feet per second airspeed error. The rather large threshold associated with perception of sink rate error was a consequence of assuming that the pilot attempts to obtain this information from the velocity of the glideslope indicator. In addition, a "residual noise" of 0.5 degree was associated with perception of pitch error to account for the lack of an explicit zero-error reference.

To simplify the analysis, the pilot was assumed to pay equal attention to glideslope, pitch, and airspeed indicators (and was assumed to obtain both displacement and rate information from all but the airspeed indicator). In addition, 34 percent of the attention was assumed "lost" in scanning. Thus, a relative attention of unity corresponded to relative attentions of 22 percent each to glideslope, pitch, and airspeed variables. As described in the literature, attentional and perceptual factors determined the observation noise variance associated with each perceptual input.[10] The remaining independent model parameters were time delay (0.29 second to account for both pilot and control-actuator delays) and motor noise covariance (-50 dB, relative to predicted control-rate variance).

Curves of predicted performance versus relative attention, generated via the optimal control model, are shown in Figure 2a. The quadratic performance index was based on assumed maximum allowable excursions for important system variables, as described earlier in this paper. Variations in "attention" were reflected by appropriate manipulation of the baseline observation noise/signal ratio as described in Levison.[2] As recommended by the military flying quality specifications,[8] the analysis was performed using a Dryden gust model having parameters appropriate to an altitude of 500 feet and longitudinal and vertical rms gust amplitudes of 10 and 6.6 feet per second, respectively.

Figure 2a shows the following trends:

1. Best achievable performance (i.e., lowest cost) with the baseline aircraft (Configuration 1) and the augmented AST (Configuration 5).
2. Worst performance, and greatest sensitivity to attentional workload, with the unstable Configurations 3 and 4.
3. Intermediate performance with the configuration having a mild instability (Configuration 8) and the vehicle having adverse dγ/dV (Configuration 2).
4. Negligible effects due to simulation of flexible modes (Configurations 6 and 7).

As noted earlier, the scope of the manned simulation study was limited to six experimental configurations. On the basis of this analysis, Configurations 4 and 5 were dropped from further consideration as they appeared to be similar in terms of performance/workload tradeoffs to Configurations 3 and 1, respectively.

Application of the rating expression of Equation (1) to the performance/workload predictions shown in Figure 2a yielded unreasonably large Cooper-Harper ratings (e.g., a rating of 8 for the baseline configuration). Partly for this reason, and partly because the -8785B backup document[11] indicates that the initial choice of gust intensities represents a low probability (1 percent) of occurrence, gust intensities were halved for subsequent analysis and experimentation. The reduced levels represent a 50-percent probability of occurrence.

In addition to the reduction in gust levels, changes in other independent model parameters were made prior to reanalysis. The allowable performance "window" for glideslope error was increased to 1 dot to reflect published Category II specifications.[12] The performance window for sink rate was increased from around 4 feet per second to 7.5 feet per second to allow for the fact that, in actual flight, the flare maneuver would substantially reduce the sink rate prior to impact. We assumed that the pilot would obtain sink rate information from the vertical speed indicator, and we decreased the perceptual threshold to 0.8 feet per second to reflect assumed visual resolution capabilities with respect to this instrument. The maximum acceptable value for the rate of thrust change was reduced to one-fifth the corresponding limit on thrust deviation to more strongly reflect the pilot's aversion to frequent changes in throttle setting. Finally, the OCM was used to predict optimal allocation of attention.

The six configurations retained for the simulation study were reanalyzed as described above; the resulting performance/workload tradeoffs are shown in Figure 2b. As is the case with the initial analysis, the penalty for relatively low attention is greatest for Configuration 3, and inclusion of flexible modes

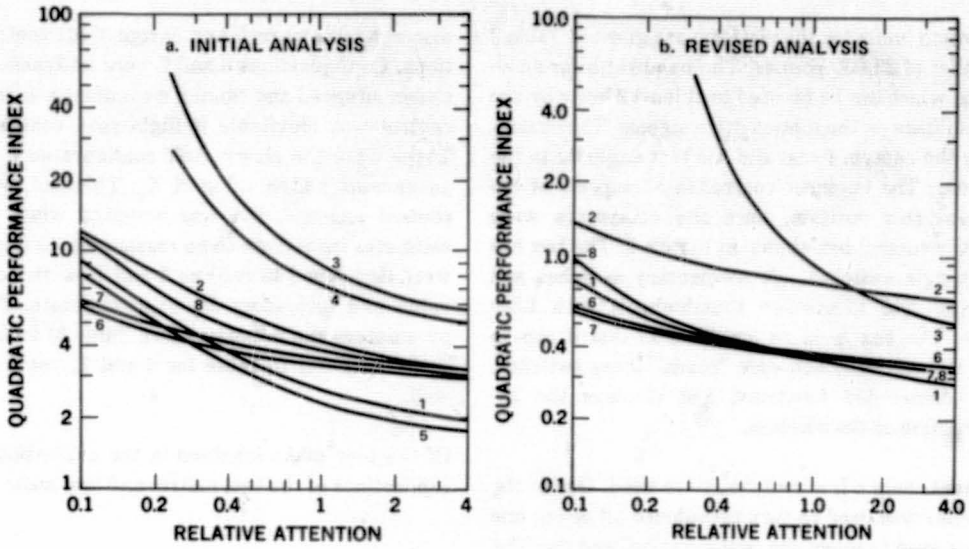


FIGURE 2. PERFORMANCE AND WORKLOAD TRADEOFFS

has little predicted influence. The predicted performance/workload tradeoff curves are compressed, however, with little separation among the curves for Configurations 1, 3, 7, and 8 at all but the lower attentional levels. Application of the rating expression of Equation (1) yields predicted ratings (shown later in this paper) that range from Level 1 to Level 3 and are consistent with those observed experimentally for Configurations 1 and 3 in the preceding study.[2]

DESCRIPTION OF EXPERIMENTS

SIMULATION CHARACTERISTICS

The simulation model used with all configurations was a complete airplane. Both longitudinal and lateral-directional degrees of freedom and controls were provided. The controls (column, wheel, and pedals) were DC-10 hardware. Control feel, force gradients, and motion limits were based on the DC-10. A full set of DC-10 instruments was provided. A flight director display was available but not used as this would affect workload and performance and, thus, pilot opinion of flying qualities. Actuator and engine dynamics typical of wide-body aircraft were simulated. Standard linearized equations of motion were used in the simulation. Euler integration of the differential equations was performed at 20 hertz. The actuators and other elements with fast dynamics were simulated using difference equations.

The simulator is Douglas' research and development motion base simulator. The motion platform is shown in Figure 3 and the cab interior in Figure 4. The cab is a DC-10 cockpit with stations for captain, first officer, flight engineer, and observer. Synthetic outside vision is available but was not used in this

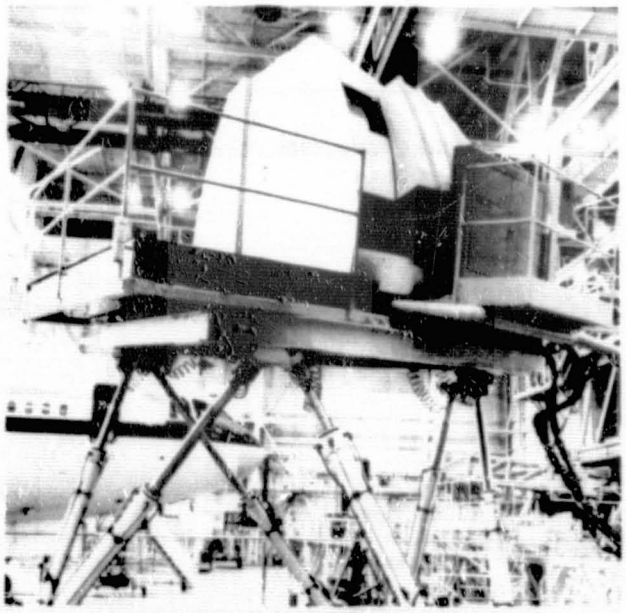


FIGURE 3. SIX-AXIS MOTION BASE SIMULATOR

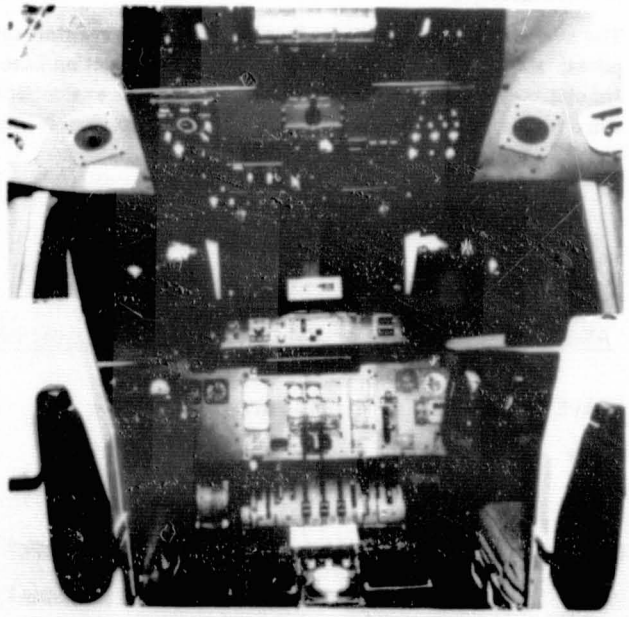


FIGURE 4. MOTION BASE SIMULATOR TRANSPORT COCKPIT

experiment. Motion limits for the platform are given in Table 2 for a moving mass of 22,000 pounds. The bandwidth for small inputs is 1 hertz, which can be boosted to at least 2 hertz by use of pre-emphasis filters on the motion drive signals. The evaluation pilot sat in the captain's seat and the test engineer in the first officer's seat. The engineer controlled all aspects of the experiment from this position, once the computers were started, using the control box shown in Figure 5. The box has six 3-position toggle switches, six momentary switches and sense lights, and five 16-position thumbwheels with LED readouts above. The box is on an umbilical so that it can be moved around the cab. The software "reads" these switches, performs the commanded functions, and displays the appropriate information on the displays.

In this experiment, only a few switches were used. One of the 3-position switches was used to turn turbulence off or on, one thumbwheel was used to select the configuration, and the pilot number was set using another thumbwheel. Three digits of the LED display showed the run number, another the pilot number, and another the configuration number currently being used by the computer. Another panel contained pushbutton switches to control start, stop, reset, and other operational functions.

The task flown was a manual instrument landing system (ILS) approach using raw rather than director, glideslope, and localizer data. Plan and side views of the approach geometry are shown in Figure 6. The Dryden turbulence model[8] was used, with 50th percentile intensities and scale lengths for an altitude of 500 feet. This model actually varies with airspeed and altitude, but was "frozen" in the experiment to match the stationary nature of the OCM. A sum-of-sines implementation was used, which concentrates the power at discrete frequencies. Twelve discrete frequencies from 0.0838 to 12.57 radians per second were used.

EVALUATIONS

The evaluations were made by four Douglas experimental test pilots, all of whom had prior experience in motion-base-simulator evaluations of flying qualities. Before the evaluations were begun, a checkout pilot flew the entire test matrix. No discrepancies were found except that, contrary to all predictions, Configurations 6 and 7 were unflyable. The coupling between airspeed and column movement was so tight that loss of control was inevitable if flight-path control was attempted. These were the elastic AST configurations. The problem was an unusually high value of X_q. This had been noticed in the pretest analysis, but was accepted when the flying-quality estimates turned out to be reasonable. Configurations 9 and 10 were developed to replace 6 and 7 by reducing X_q to a small value and increasing X_u to compensate. The flying quality parameters and estimates (see Table 3) for 9 and 10 were virtually identical to those for 6 and 7, yet 9 and 10 flew quite well.

Of the four pilots involved in the evaluations, two made four replications of the test matrix and two made five. Each session

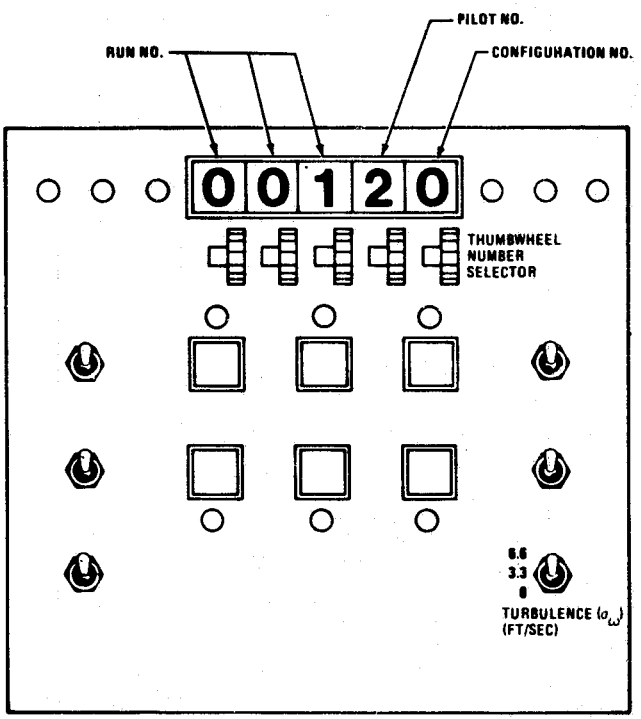


FIGURE 5. SIMULATOR CONTROL BOX

TABLE 2
MOTION LIMITS

AXIS	EXCURSION	VELOCITY	ACCELERATION
HEAVE	±42 IN.	39 IN./SEC	1.65 g
SWAY	±67.5 IN.	67 IN./SEC	2.43 g
SURGE	±65 IN.	71 IN./SEC	1.50 g
ROLL	±30.7 DEG	35.6 DEG/SEC	7.8 RAD/SEC^2
PITCH	±33.3 DEG	33.6 DEG/SEC	7.8 RAD/SEC^2
YAW	±38.7 DEG	36.3 DEG/SEC	7.9 RAD/SEC^2

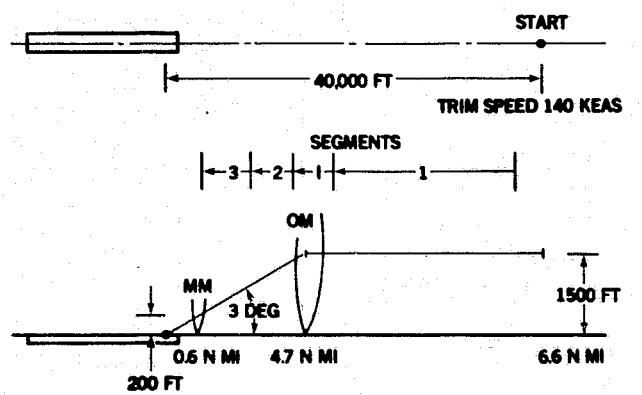


FIGURE 6. APPROACH GEOMETRY

began with a briefing in which the test procedure and performance standards were reviewed. The pilot then flew several approaches to warm up, or get used to, the equipment and procedure. He then flew two approaches with each configuration (turbulence off, then on). The configurations were presented in pseudo-random order, with the order balanced across replications. The "turbulence off" runs were flown to allow additional practice, to isolate the turbulence effects, and to gather data for the development of a glideslope capture strategy. The pilots could do whatever maneuvers and experiments they wanted in these runs after the intercept maneuver. In the "turbulence on" runs, however, they were told to track the ILS to the performance standards at all times. A replication of the test matrix took 1-1/2 to 2 hours. A total of 319 approaches were flown for data in the test; 25 more were flown by the checkout pilot. At the end of the test, the pilots were quizzed again about a number of items, including the performance standards they flew to in the test and how they allocated their attention.

A great deal of objective and subjective data was taken. Time histories of 50 variables were recorded on digital magnetic tape at 5 hertz for every approach. Stripchart records of 16 variables were made. The mean, root mean square, maximum and minimum values, and standard deviation of 15 variables were computed on-line and output on a line printer at the end of each run. Instantaneous values of 10 variables at 10 points along the approach were also printed out. The line printer was also used to record bookkeeping information, such as run start time and date, run number, configuration number, pilot number, etc., to reduce the test engineer's workload. The subjective data taken included Cooper-Harper pilot ratings, effort ratings for the three subtasks and three aspects of control, and pilot comments. The engineer made brief handwritten notes to supplement the complete record made by the cockpit voice recorder.

EXPERIMENTAL RESULTS

Statistical analysis was performed on both Cooper-Harper ratings and closed-loop performance metrics. Ratings were first averaged across replications to obtain an average rating per condition per pilot. Population means and across-subject standard deviations were then computed from the individual subject means.

Statistics on system "errors" (deviations from trim) were computed for the three steady-state-like segments of the approach. Results for the final segment of the glideslope tracking task — corresponding to descent from approximately 700 to 200 feet altitude — are reported here. The mean and variability components of each error variable were analyzed separately. Only response variability is reported here, as only that error component can be compared with model predictions for the glideslope tracking task (remember that the external disturbances were zero-mean processes). Mean error is primarily reflective of piloting strategy (e.g., carry excess airspeed, "duck under" the glideslope) and therefore is not treated directly by the OCM. In general, the variability component was dominant. A variance score was first computed for each error variable of interest within a given replication. The square root of this measure was then treated as the basic error score. Note that this measure reflects within-trial variability, not run-to-run or pilot-to-pilot variability. Error scores computed in this manner were then subjected to the same statistical analysis as that described earlier for the pilot ratings.

TABLE 3
FLYING QUALITY LEVELS OF TEST CONFIGURATIONS 1 THROUGH 10

CONFIG NO.	$\omega_{n_{sp}}$ vs n/α	ζ_{sp}	ζ_{ph} or $T_{2_{ph}}$	STATIC STABILITY	dγ/dV	-8785 OVERALL	BAND-WIDTH	BW + dγ/dV
1	1	1	1	STABLE	1	1	1	1
2	1-1/2	1	1	STABLE	WORSE THAN 3	WORSE THAN 3	1	2-1/2
3	WORSE THAN 3	2	WORSE THAN 3	UNSTABLE	1	WORSE THAN 3	3	3
4	WORSE THAN 3	1	WORSE THAN 3	UNSTABLE	3	WORSE THAN 3	3	3
5	2	1	1	STABLE	1	2	1	1
6	3-1/2	2	1	STABLE	1	3-1/2	1	1
7	3	1	1	STABLE	1	3	1	1
8	WORSE THAN 3	1	WORSE THAN 3	UNSTABLE	1	WORSE THAN 3	2	2
9	3	2	1	STABLE	1	3	1	1
10	2	1	1	STABLE	1	2	1	1

Measured and predicted pilot opinion ratings are presented in Figure 7a. Across-subject standard deviations (designated by brackets) were generally less than one rating unit. Thus, the experimental technique yielded rating predictions that were reasonably consistent across pilots. The trend of the experimental ratings agreed well with pre-experimental model predictions: Configurations 1, 8, 9, and 10 were rated similarly whereas Configurations 2 and 3 received ratings that were appreciably more adverse. The major discrepancy between prediction and experiment was the relative compression of the simulation results, with the "better" configurations receiving Level 2 rather than the predicted Level 1 ratings. In addition, Configurations 2 and 3 were rated nearly the same on the average, whereas the analytic technique predicted a 2-unit spread.

In a previous study in which lateral-directional characteristics were considered to reflect Level 1 handling qualities, the "baseline" Configuration 1 received an average pilot rating in the Level 1 range.[4] Lateral characteristics were less favorable for the study reported here, receiving rating scores in the Level 2 range. Thus, we suspect that the greater-than-expected rating scores for Configurations 1, 8, 9, and 10 reflected, in part, an interaction with the lateral-axis tasks. (Model predictions were based on the assumption that the lateral-axis task would present no appreciable handling-quality problems.)

Configuration 2 was also explored in the previous study. In that study, as well as in the current one, the rating score obtained in the simulation study was higher than predicted analytically. As discussed shortly, this model/experiment difference may be due in part to a failure of the analytic scheme, as described so far, to consider the adverse effects of requiring loop closures that are not part of the pilot's standard repertoire.

Predicted and experimental measures of the quadratic performance index are compared in Figure 7b. Two sets of model predictions are shown: scores obtained with relative attentions corresponding to minimum ratings as determined from the expression of Equation (1), and scores corresponding to a relative attention of unity. Although measured scores were considerably greater than predictions, predicted trends were confirmed. As with the rating scores, performance scores for Configurations 1, 8, 9, and 10 were similar, whereas substantially greater (less favorable) scores were observed for Configurations 3 and 4.

A comparison of predicted and measured "error" variability scores for selected response variables is given in Figure 8. Again, measured scores were greater than analytic predictions, but trends related to the effects of vehicle characteristics were generally in agreement. In particular, the analytic procedure correctly predicted that relatively large elevator deflections would be required for Configuration 3, whereas large thrust changes would be required for Configuration 2. Overall, the two methods of predicting objective performance scores replicated experimental trends with similar fidelity.

Additional model analysis was conducted to determine methods for obtaining a more accurate assessment of the adverse handling qualities associated with Configuration 2, and for predicting the severe controllability problems found experimentally with Configurations 6 and 7. Compared to the baseline configuration, these three required a strategy that relied more heavily on throttle for height control and elevator for speed control: Configuration 2 because of adverse dγ/dV characteristics, and Configurations 6 and 7 because of a high pitch/speed coupling. This observation suggested a simple technique for analytically detecting handling quality problems associated with undesirable throttle activity: model analysis was performed with and without the throttle control active. To test the discrimination of the procedure, model analysis was

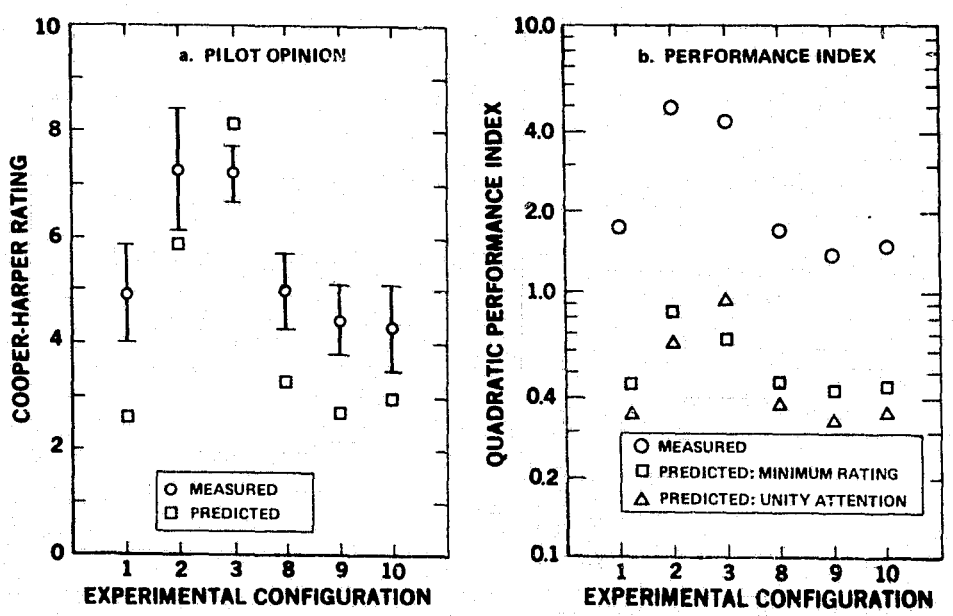


FIGURE 7. COMPARISON OF CRITERIA PREDICTIONS AND RESULTS

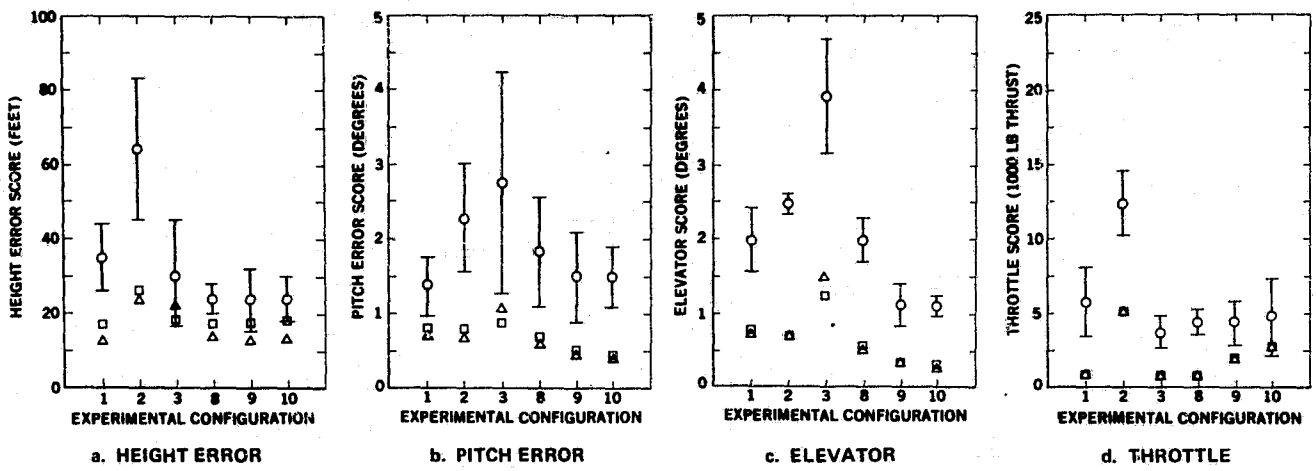


FIGURE 8. PREDICTED AND MEASURED PERFORMANCE AND WORKLOAD SCORES

performed again for Configuration 1 (baseline), Configuration 3 (greatest instability), and Configuration 9 (configuration 6 modified to reduce the pitch-speed coupling). This analysis was performed with the baseline observation noise/signal ratio adjusted to reflect unity relative attention.

Figure 9 shows that this method readily identified handling quality problems related to throttle activity. The predicted quadratic performance indices for Configurations 1, 3, and 9, while different from each other, were relatively unaffected by the exclusion of throttle control. On the other hand, omission of throttle control caused the performance metric to more than double for Configuration 2 and to increase nearly sevenfold for Configuration 6. Thus, a model comparison of this sort appears to be a simple device for predicting handling quality difficulties caused by requirements for significant throttle activity.

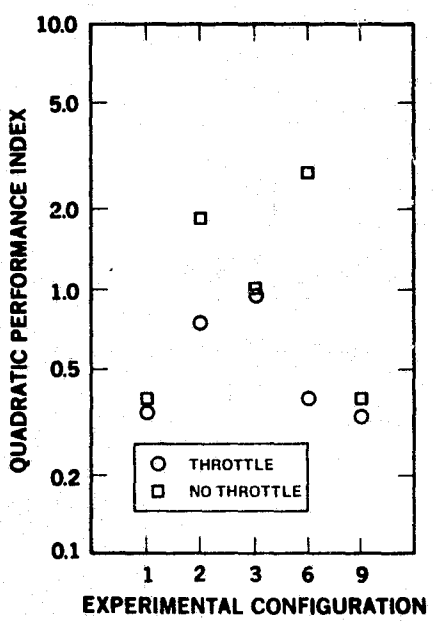


FIGURE 9. EFFECT OF THROTTLE CONTROL ON PREDICTED PERFORMANCE INDEX

CONCLUSIONS

A number of conclusions can be drawn from the analytical and experimental results of the subject study program.

1. A closed-loop criterion, or estimator of flying qualities, has been developed and validated against experimental data. The development of the model comprising the criterion was based on the characteristics of the task being modeled, not on the data. That is, it is not simply a model fit to a set of data, it is a simulation model. The characteristics used to develop the model include pilot preferences (determined from interviews), human capabilities (determined from laboratory experiments), the physics of the situations, and engineering judgment.
2. In the experiment, repeatable Cooper-Harper pilot opinion ratings were obtained by strict experiment protocol. Important aspects of the experiment design were:
   a. Design of a task with only zero-mean, stationary, random disturbances
   b. Well-defined subtasks and associated performance standards
   c. Adequate pilot familiarization.
3. A data base has been developed and recorded which includes pilot ratings and objective performance measures. The data have been used in validating the optimal control model as described herein, and will be used for further development in the future.
4. The model correctly predicted an absence of an effect of the structural modes simulated on handling qualities.
5. The model correctly predicted the experimental trends in workload, performance, and pilot ratings. This was an actual prediction: it was done before the experiment.
6. The ability of the model to predict low-frequency flight path control problems was demonstrated. These problems are detected using a two-step process. First, model predictions are made assuming throttle as a control. Then model predictions are made without throttle as a control

and compared with the previous analysis. If there is a major degradation, flight path control problems can be expected.

7. The experimental error scores were about twice those predicted analytically. The potential causes of this discrepancy include:
   a. Perceptual and indifference thresholds were perhaps greater in a realistic flight task than in a laboratory tracking task
   b. Possible discontinuous control behavior by test pilots
   c. Interference between longitudinal and lateral axes
   d. Less training in this experiment than in many laboratory tracking experiments.

   The discrepancies are not readily attributable to differences between predicted and actual tradeoffs of error and control: path, altitude, speed, and control variations were all greater than predicted.
8. Given the state of the art, the analytic scheme developed here is recommended for use in predicting important trends rather than absolute performance scores.
9. It is not necessary to select attention according to the rating expression, Equation (1). The quadratic performance index predicted for constant attentional workload mimics trends of pilot ratings. These results confirm an earlier study of Hess.[1]

REFERENCES

1. Hess, R.A., "Prediction of Pilot Opinion Ratings Using an Optimal Pilot Model", Human Factors, 1977, 19(5).
2. Levison, W.H., "A Model-Based Technique for Predicting Pilot Ratings for Large Commercial Transports", NASA CR-3257, April 1980.
3. Levison, W.H., "A Model-Based Technique for Predicting Pilot Opinion Ratings for Large Commercial Transports", Proc. of the 16th Annual Conference on Manual Control, Cambridge, MA, May 1980.
4. Rickard, W.W., "Longitudinal Flying Qualities in the Landing Approach", Proc. of the 12th Annual Conference on Manual Control, NASA TM X-73, 170, May 1976.
5. Cooper, G.E., and Harper, R.P., Jr., "The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualitites", NASA TN D-5153, April 1969.
6. Levison, W.H., Elkind, J.I., and Ward, J.L., "Studies of Multi-Variable Manual Control Systems: A Model for Task Interference", NASA CR-1746, May 1971.
7. Levison, W.H., "A Model for Mental Workload in Tasks Requiring Continuous Information Processing", Mental Workload Its Theory and Measurement, Plenum Press, New York and London, 1979.
8. "Flying Qualities of Piloted Airplanes", Military Specification, MIL-F-8785B (ASG), 1969.
9. Levison, W.H., "The Effects of Display Gain and Signal Sandwidth on Human Controller Remnant", AMRL-TR-70-93, March 1971.
10. Baron, S., and Levison, W.H., "The Optimal Control Model: Status and Future Directions", Proc. of the 1980 International Conference on Cybernetics and Society, Cambridge, MA, October 1980.
11. Chalk, C.R., Neal, T.P., Harris, T.M., Pritchard, F.E., "Background Information and User Guide for MIL-F-8785B (ASG), 'Military Specification-Flying Qualtities of Piloted Airplanes' ", AFFDL-TR-69-72, August 1969.
12. "DC-10 Flight Study Guide", Douglas Aircraft Company, June 1975.

AN ANALYTICAL APPROACH FOR PREDICTING PILOT INDUCED OSCILLATIONS

By Ronald A. Hess

NASA Ames Research Center

SUMMARY

The optimal control model (OCM) of the human pilot is applied to the study of aircraft handling qualities. Attention is focused primarily on longitudinal tasks. The modeling technique differs from previous applications of the OCM in that considerable effort is expended in simplifying the pilot/vehicle analysis. After briefly reviewing the OCM, a technique for modeling the pilot controlling higher order systems is introduced. Following this, a simple criterion for determining the susceptability of an aircraft to pilot-induced oscillations (PIO) is formulated. Finally, a model-based metric for pilot rating prediction is discussed. The resulting modeling procedure provides a relatively simple, yet unified approach to the study of a variety of handling qualities problems.

INTRODUCTION

The advent of modern, digital stability and control augmentation systems has created a renewed interest in the study of aircraft longitudinal handling qualities. This renewed interest is attributable to two factors: First, the higher order nature of the dynamics typically associated with digital control systems make analytical prediction of handling qualities difficult. Contemporary handling qualities specifications[1] are written assuming "classical" aircraft characteristics, e.g., in the longitudinal mode, the existence of distinct and dominant short-period dynamics is assumed. With modern systems, the short-period characteristics may be dramatically altered by feedback and the higher order control system dynamics, themselves, may dominate the vehicle handling qualities. Second, shortcomings in predictive techniques are made even more critical by the fact that severe handling qualities deficiencies often arise in practice which are directly attributable to the higher order nature of the digital control law implementation. An example of this is the ability of high frequency phase lags or time delays in the control system to sharply degrade aircraft handling qualities and to be a contributing factor to pilot induced oscillations.[2]

In the research to be described, a pilot modeling technique for handling qualities research, discussed in Ref. 3, is utilized and extended to cover higher order systems. The characteristics of over thirty aircraft configurations are analyzed, primarily in the longitudinal mode. Particular emphasis is placed upon those configurations where control system dynamics and time

delays have been recognized as contributing factors to handling qualities dificiencies. The contribution of vehicle/control system dynamics to PIO tendencies are outlined and a metric for pilot rating prediction is discussed.

BACKGROUND

The pilot modeling technique as discussed in Ref. 3 forms the framework for the research described here. This technique utilizes the optimal control model of the human pilot and a novel method for the a priori selection of dominant OCM parameters (index of performance weighting coefficients and observation noise/signal ratios). A brief tutorial review of the procedure for selecting index of performance weighting coefficients is now presented. Consider the longitudinal tracking task of Fig. 1 in which the pilot is attempting to minimize pitch attitude deviations $\theta(t)$ in the presence of atmospheric disturbances. Ignore the dashed "internal attitude command" for the present. An acceptable index of performance for this task would be[3]

$$J = E\left\{\lim_{X\to\infty}\frac{1}{2X}\int_{-X}^{X}[\theta^2(t)/\theta_M^2 + \dot{\delta}^2(t)/\dot{\delta}_M^2]dt\right\} .$$

where $\dot{\delta}(t)$ is control rate.

As discussed in Ref. 3, we assign an arbitrary maximum allowable deviation to the time rate of change of the error, $\dot{\theta}(t)$, and denote it $\dot{\theta}_M$. Now an effective time constant T can be introduced to define maximum allowable deviations of the integral and derivatives of $\dot{\theta}(t)$ as:

$$\theta_M = \dot{\theta}_M T \ ;$$

$$\dot{\theta}_M = \text{specified but arbitrary} \ ;$$

$$\ddot{\theta}_M = \dot{\theta}_M/T \ ; \qquad (1)$$

and

$$\dddot{\theta}_M = \ddot{\theta}_M/T = \dot{\theta}_M/T^2 \ .$$

The justification for using a single time constant to represent the ratio of the maximum value of a variable to that of its next highest derivative rests upon the system bandwidth implications which follow when Eqs. (1) are used in implementing the OCM. We will also assign a maximum allowable deviation to the time rate of change of the pilot's control, $\dot{\delta}(t)$, and denote it $\dot{\delta}_M$. Similar to Eqs. (1) we write

$$\delta_M = \dot{\delta}_M T \quad ;$$

$$\dot{\delta}_M = \text{to be selected} \quad ;$$

$$\ddot{\delta}_M = \dot{\delta}_M / T \quad ;$$

$$\dddot{\delta}_M = \ddot{\delta}_M / T = \dot{\delta}_M / T^2 \quad ; \tag{2}$$

.

.

.

The value of $\dot{\delta}_M$ is not arbitrary, however, but is found using Eqs. (1) and (2) and the vehicle dynamics as follows: Let the pitch attitude dynamics of the the aircraft be given by

$$\frac{\theta}{\delta}(s) = K \frac{s^{n-1} + a_{n-2} s^{n-2} + \cdots + a_1 s + a_0}{s^n + b_{n-1} s^{n-1} + \cdots b_1 s + b_0} \quad . \tag{3}$$

Then, as explained in Ref. 3, we write

$$\dot{\delta}_M = \frac{[1/T^{n-1} + |b_{n-1}|/T^{n-2} + \ldots + |b_1| + |b_o|T]}{K[1/T^{n-2} + |a_{n-2}|/T^{n-3} + \ldots + |a_1| + |a_o|T]} \dot{\theta}_M \quad . \tag{4}$$

Thus, once T is known, $\dot{\delta}_M$ and θ_M (and, if needed, $\ddot{\theta}_M$, etc.) can be determined immediately. Choosing T involves selecting a domain of 1/T: $1/4\tau < 1/T < 4/\tau$ and then plotting J, the value of the OCM index of performance, vs 1/T. The operating point or "knee" of this curve determines T. The knee is defined as the point where

$$\frac{\partial J}{\partial \log(1/T)} = \eta_\delta \frac{J\big|_{T=\tau/4} - J\big|_{T=4\tau}}{\log(4/\tau) - \log(1/4\tau)} \quad . \tag{5}$$

Here η_δ is a constant, nominally unity, which can be used to reflect manipulator characteristics, much like an efficiency factor; τ is the pilot's time delay (nominally 0.2 secs). $J\big|_{T=\tau/4}$ is the value of the index of performance which results when $T=\tau/4$.

The ability of the OCM parameter selection technique to provide a pilot model which matches measured pilot describing fucntions, remnant power spectral densities and root mean square (RMS) performance measures was demonstrated in Ref. 3. In addition, the modeling technique was shown capable of providing qualitative and quantitative handling qualities assessments. The method for

selecting observation noise/signal ratios for the OCM is discussed in Ref. 3 and will not be dealt with here.

Although Eq. (3) shows dynamics of arbitrary order, all the pitch attitude dynamics of Ref. 3 were of the form:

$$\frac{\theta}{\delta}(s) = \frac{K_\theta(s+1/T_L)}{s(s^2+2\zeta_n\omega_n s+\omega_n^2)} \quad . \tag{6}$$

When higher order dynamics are encountered, the method for selecting the operating point needs to be modified slightly. The large phase lags typically associated with the dynamics of vehicles with higher order dynamics need to be reflected in choosing the domain of 1/T to be used in Eq. (5). To accomplish this, a delay τ_D is defined as the delay which accrues when the vehicle dynamics of Eq. (3) are represented as

$$\frac{\theta}{\delta}(s) = \frac{K_\theta(s+1/T_L)e^{-\tau_D s}}{s(s^2+2\zeta_n\omega_n s+\omega_n^2)} \quad . \tag{7}$$

The parameters on the right hand side of Eq. (7) are found using a program to fit a linear transfer function model to the actual vehicle dynamics.[4] Equation (5) is modified by simply replacing τ with $\tau+\tau_D$. The resulting equation is interpreted graphically in Fig. 2. Calculating τ_D and including it in Eq. (5) constitutes the extension of the methods of Ref. 3 to higher order systems. It is important to emphasize that the actual higher order vehicle dynamics are used in the modeling procedure; Eq. (7) is employed only to select τ_D which, in turn, determines the domain of 1/T used in finding the index of performance weighting coefficients.

APPLICATION TO AIRCRAFT LONGITUDINAL HANDLING QUALITIES

Pilot Induced Oscillations

Table I lists the aircraft configurations which have been analyzed in this study. The designations in the column labeled "Configuration" use notation found in the corresponding references. The first sixteen deal with high performance fighter-type aircraft in tracking or landing approach conditions and are taken from Refs. 2, 5 and 6. These configurations constitute the test cases for the majority of the assessments. The next four configurations are taken from Ref. 7 and represent pilot-in-the-loop simulations of a hovering helicopter. Configurations 21-25 are flight test results from Ref. 9 in which the Princeton University Variable Response Aircraft (VRA) was used to determine the effect of digital sampling rates and time delays on longitudinal handling qualities. The vehicle dynamics appropriate for 105 kts airspeed were used in the modeling procedure. The pilot ratings which were used were average values obtained from altitude tracking and approach and landing tasks (Fig. 3 of Ref. 9). Finally, configurations 26-32 are taken from Ref. 10 where a moving base simulator experiment on the NASA Ames Flight Simulator for Advanced

Aircraft (FSAA) was described which investigated a wings level-turn control mode for air-to-ground weapons delivery. Note that unlike the previous twenty five configurations, these involve lateral-directional aircraft handling qualities. The effective vehicle dynamics for the lateral gunsight aiming task were parameterized by a damping ratio ζ_n, an undamped natural frequency ω_n and a pure time delay τ_D.[10] The data for the so-called "fine" task were used. This task is explained in Ref. 10.

As an example of the modeling results, Fig. 3 shows the longitudinal open-loop pilot/vehicle characteristics (Y_pY_c) for three of the configurations used in Ref. 2. Here, the NASA Dryden F-8 digital fly-by-wire aircraft is considered with a rudimentary augmentation system ("Pitch Direct") and three transport time delays of 0.13 sec, 0.23 sec, and 0.33 sec, respectively. The predicted effect of the time delays is apparent in the reduced open-loop crossover frequencies ω_c. This open-loop characteristic obviously has a deleterious effect on the closed loop θ/θ_c transfer functions as shown in Fig. 4 ($\theta/\theta_c=Y_pY_c/(1+Y_pY_c)$). This transfer function is important is assessing PIO susceptability. Although the task has been defined as pitch-attittude disturbance regulation, attitude commands θ_c internally generated by the pilot would be employed in precise altitude regulation (dashed line in Fig. 1). Note in Fig. 4, that as τ_D increases, $|\theta/\theta_c|$ and $\angle\theta/\theta_c$ decrease at all frequencies. Perfect command following, of course, implies $\theta/\theta_c=1.0$ at all frequencies. In Fig. 4, $|\theta/\theta_c|<1.0$ for all configurations when $\omega<3.0$ rad/sec, and is particularly poor for the configuration with $\tau_D=0.33$ sec. It can be readily shown that open-loop crossover frequencies less than 3-4 rad/sec will invariably result in poor closed-loop attitude command-following characteristics. The simplest and most direct way for the pilot to attempt to improve this closed loop command-following performance is to increase ω_c by increasing his static gain. If the pilot attempts this for the F-8 configuration with $\tau_D=0.33$ sec, a very lightly damped closed-loop oscillation occurs at $\omega=3.3$ rad/sec (see Fig. 4). This is identical to the PIO frequency shown in Ref. 2 for this configuration.

Similar results are also obtained for configurations from Ref. 5. Figure 5 compares a pair of open-loop transfer functions obtained using configurations "11" and "12" from Ref. 5 and applying the pilot modeling technique discussed above. Once again, the dramatic difference in the crossover frequencies ω_c is apparent. The effects of the pilot's attempting to improve the performance of configuration "12" by increasing his static gain by 10 dB are shown in Fig. 6. Once again, a lightly damped oscillatory mode is seen to appear. The simulations of Ref. 5 were intended to provided performance comparisons for configurations which were flight tested and discussed in Ref. 8. The latter report included Pilot-Induced-Oscillation-Ratings (PIOR) obtained using the scale of Fig. 7. It is interesting to note that configuration "11" received an average PIOR or 1 indicating a very satisfactory vehicle whereas configuration "12" received a marginal average rating of 2.7 indicating a vehicle with definite PIO tendencies. These experimental results are seen to corroborate the analytical findings just discussed.

Next, consider two configurations from Ref. 6 denoted as "4-1" and "6-1". Figure 8 shows the Y_pY_c plots for these configurations. Configuration "4-1"

received a very satisfactory PIOR of 1 whereas configuration "6-1" received a very poor PIOR of 4. Indeed, configuration "6-1" produced a PIO in flight test with a frequency of approximately 3.75 rad/sed. Analytically increasing the pilot's static gain by 4.75 dB (the limit for closed-loop stability) in the modeling-results for this configuration produced a closed-loop oscillation at approximately 3.5 rad/sec. This 4.75 dB increase would increase ω_c from around 1.5 rad/sec to only around 2.5 rad/sec as compared to a value of 4.5 rad/sec for configuration "4-1".

Figure 9 shows the predicted Y_pY_c's for a pair of configurations from Ref. 9. The task was longitudinal control in approach and landing using the Princeton VRA. The variable of interest here was the amount of effective delay in the control system. In the first, an effective delay of 0.055 sec was employed, while in the second, 0.355 sec was used. Again, note the striking difference in crossover frequencies in the predicted pilot/vehicle dynamics. In the first case, ω_c=3.4 rad/sec, while in the latter, ω_c=0.55 rad/sec. Flight test of the first configuration showed no PIO tendencies, while those for the latter produced PIO's.[9]

Finally, Fig. 10 shows the predicted Y_pY_c's for a pair of configurations from Ref. 10. In the first, the control system parameters were ζ_n=1.4, ω_n= 2.0 rad/sec and τ_D=0.0 sec, while in the second, ζ_n=1.4, ω_n=15.0 rad/sec and τ_D=0.49 sec. The ω_c difference is again evident. Simulation results indicated that the configuration with delay was definitely PIO prone and the one without delay was not. It is interesting to point out that the configuration without delay still received an average Cooper-Harper pilot rating of 6.5, even though it was not PIO prone. Thus, poor pilot ratings, per se, are not a necessary condition for PIO susceptability.

In each of the cases above, we have made direct comparisons of vehicles which were found to be PIO prone with those which were not. This was done to emphasize the fact that the method proposed here is clearly discriminatory in predicting PIO susceptability. The simple criterion for exonerating a vehicle from PIO tendencies requires that the predicted pilot/vehicle crossover frequencies associated with inner attitude-loops be greater than 3-4 rad/sec.

Cooper-Harper Ratings

Figure 11 is a plot of the Cooper-Harper ratings which the thirty-one configurations from Table I received in simulation or flight test vs the value of a proposed handling qualities metric defined as $K_i \cdot ((\tau+\tau_D)/\tau)^4 \cdot J$. No ratings were reported in Ref. 10 for configuration 32 of Table I. Hence, only thirty one data points are shown in Fig. 11. The K_i can be interpreted as a "calibration parameter" which, when multiplied by $((\tau+\tau_D)/\tau)^4 \cdot J$, allows the reported pilot ratings from different tasks and data sources to coalesce as shown in Fig. 11. Note that we do <u>not</u> allow K_i to vary within the analysis of any particular task, regardless of configuration changes. Thus, the analysis of the six configurations from Ref. 2 used a single value of K_i (call it K_1). The analysis of the seven configurations from Refs. 2 and 8 used a single value

of K_i (call it K_2), etc. In all, six different K_i values (each one corresponding to the six different symbols in Fig. 11) were used to generate Fig. 11. With the exception of K_i, all the parameters of the metric are an intrinsic part of the modeling procedure, and, as such, involve no guesswork on the part of the analyst. In order to determine K_i, the analyst must have an actual pilot rating for one of the configurations tested for the task under study. If the analyst does not have such a rating available, Fig. 11 is still useful, since the curve is nearly linear from a pilot rating of about 2.0 to 10.0, a range which covers 80% fo the Cooper-Harper scale. Thus, relative rating changes may be able to be predicted using the linear portion of the curve. Note that, with the exception of one data point (Config. 19 from Ref. 7), the scatter in the ratings in Fig. 11 is only about ±1/2 a pilot rating.

The inclusion of the factor $((\tau+\tau_D)/\tau)^4$ in the metric deserves a brief discussion. In previous research with the OCM, the value of J, alone, has been found to correlate well with pilot opinion rating.[11] In many of the configurations studied here, however (those with $\tau_D>0$), the value of J was not acceptable as a metric. In general, the "predicted" opinion rating increments were smaller than those reported in experiment. There appears to be a reason for this based upon pilot tracking performance. Namely, when the task is disturbance regulation involving relatively low-bandwidth turbulence, large time delays are not necessarily a harbinger of dramatic deterioration in tracking performance. This is analytically verified by considering the RMS tracking scores for configurstions 1 and 3 from Table I. Here, a 154% increase in time delay between configurations 1 and 3 invloves a $\log\omega_c$ regression of nearly a decade. However, the predicted RMS pitch attitude score increases by only 36% and the predicted RMS control-rate score actually decreases. As we have attempted to point out here, however, the same cannot be said for discrete command following or abrupt maneuvers. In this case, ω_c regression can have a significant impact on the ability of the closed-loop pilot/vehicle system to follow abrupt, internally generated commands. It certainly is not unreasonable to postulate that such short-term response characteristics (in addition to RMS characteristics) are reflected in pilot opinion rating. Indeed, recorded pilot comments support this notion (e.g., Refs. 2 and 6). The inclusion of $((\tau+\tau_D)/\tau)^4$ in the metric appears to account for the influence of these delays on pilot opinion in a straightforward manner, employing an easily identifiable parameter (τ_D).

CONCLUSIONS

The research summarized in this paper provides a unified approach to pilot/vehicle analysis, and in particular for

1) modeling the pilot controlling higher order systems
2) predicting the susceptability of aircraft to longitudinal PIO's
3) predicting pilot ratings for tasks when one configuration rating is known, or predicting relative rating changes between configurations.

Although the majority of tasks studied dealt with longitudinal control, five lateral-directional configurations were successfully analyzed with no changes in the modeling technique.

REFERENCES

1. Anon., "Military Specification, Flying Qualities of Piloted Airplanes: MIL-F-8785B(ASG)," August 1969.

2. Berry, D. T., Powers, Bruce G., Szalai, K. J., and Wilson, R. J., "A Summary of an In-Flight Evaluation of Control System Pure Time Delays During Landing Using the F-8 DFBW Airplane," AIAA Paper No. 80-1626, 1980 AIAA Atmospheric Flight Mechanics Confernece, Danvers, Mass.

3. Hess, R. A., "A Pilot Modeling Technique for Handling Qualities Research," AIAA Paper No. 80-1624, 1980 AIAA Atmospheric Flight Mechanics Conference, Danvers, Mass.

4. Seidel, R. C., "Transfer-Function-Parameter Estimation From Frequency Response Data - A FORTRAN Program," NASA TM X-3286, 1975.

5. Arnold, J. D., "An Improved Method of Predicting Aircraft Longitudinal Handling Qualities Based on the Minimum Pilot Rating Concept," Air Force Institute of Technology, GGC/MA/73-122, 1973.

6. Smith, Rogers, E., "Effects of Control System Dynamics on Fighter Approach and Landing Longitudinal Flying Qualities, Vol. I.," Air Force Flight Dynamics Laboratory, AFFDL-TR-73-122, 1978.

7. Miller, D. P., and Vinje, E. W., "Fixed-Base Flight Simulator Studies of VTOL Aircraft Handling Qualities in Hovering and Low-Speed Flight," Air Force Flight Dynamics Laboratory, AFFDL-TR-67-152, 1968.

8. Neal, Peter T., and Smith Rogers E., "An In-Flight Investigation to Develop Control System Design Criteria for Fighter Airplanes," Air Force Flight Dynamics Laboratory, AFFDL-TR-70-74.

9. Stengel, R. F., and Miller, G. E., "Flight Tests of a Microprocessor Control System," Journal of Guidance and Control, Vol. 3, No. 6, Nov.-Dec. 1980, pp. 494-500.

10. Sammonds, R. I., and Bunnell, J. W., Jr., "Flying Qualities Criteria for Wings-Level-Turn Maneuvering During an Air-to-Ground Weapon Delivery Task," AIAA Paper No. 80-1628, 1980 AIAA Atmospheric Flight Mechanics Conference, Aug. 11-13, 1980.

11. Hess, R. A., "Prediction of Pilot Opinion Ratings Using an Optimal Pilot Model, Human Factors, Vol. 19, No. 5, Oct. 1977, pp. 459-475.

Table I. Aircraft Configurations Analyzed

No.	Configuration				Reference
1	F-8 "Pitch Direct"	0.13 sec delay			2
2		.23			
3		.33			
4	"ISAS"	0.13 sec delay			
5		.23			
6		.33			
7	"2D"				5,8
8	"5A"				
9	"8A"				
10	"9"				
11	"10"				
12	"11"				
13	"12"				
14	"3-1"				6
15	"4-1"				
16	"6-1"				
17	"PH-28"				7
18	"PH-29"				
19	"PH-32"				
20	"PH-35"				
21	Princeton VRA	0.055 sec delay			9
22		.135			
23		.255			
24		.355			
25		.455			
	FSAA Wings-Level Turn (lateral-directional)	ζ_n	ω_n (rad/sec)	τ_D (sec)	10
26		1.4	15.0	0	
27		1.4	2.0	0	
28		2.0	8.0	0	
29		0.7	6.0	0	
30		0.5	4.5	0	
31		0.3	4.5	0	
32		1.4	4.5	0.49	

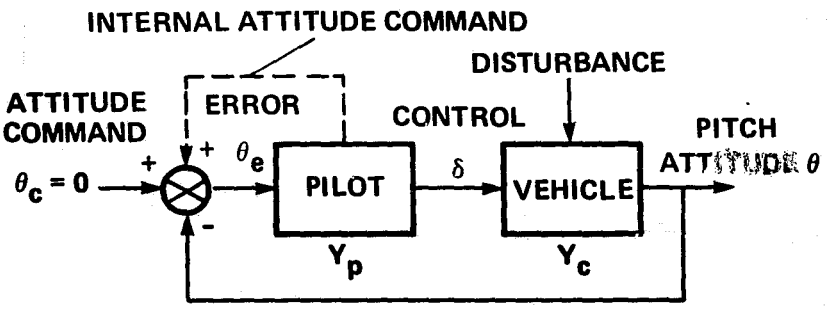


Figure 1. A Pitch Attitude Regulation Task

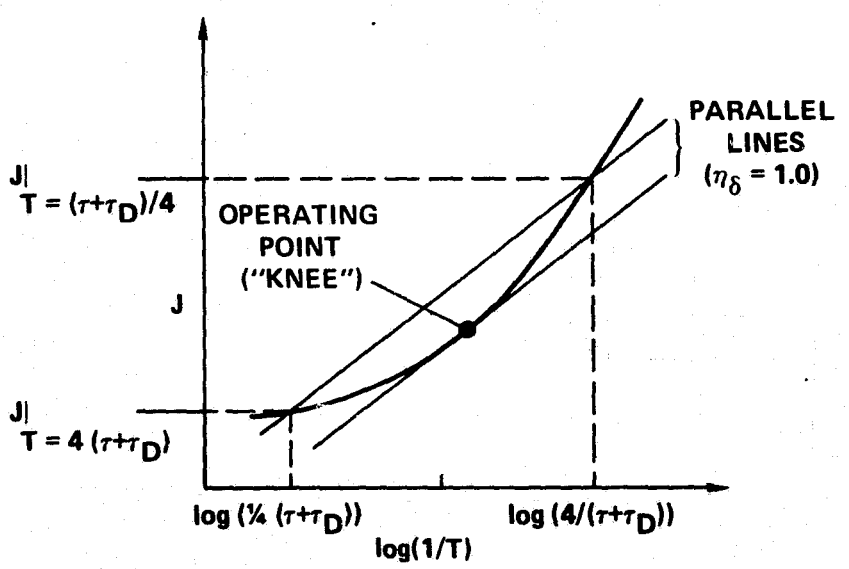


Figure 2. Selecting an "Effective Time Constant" T

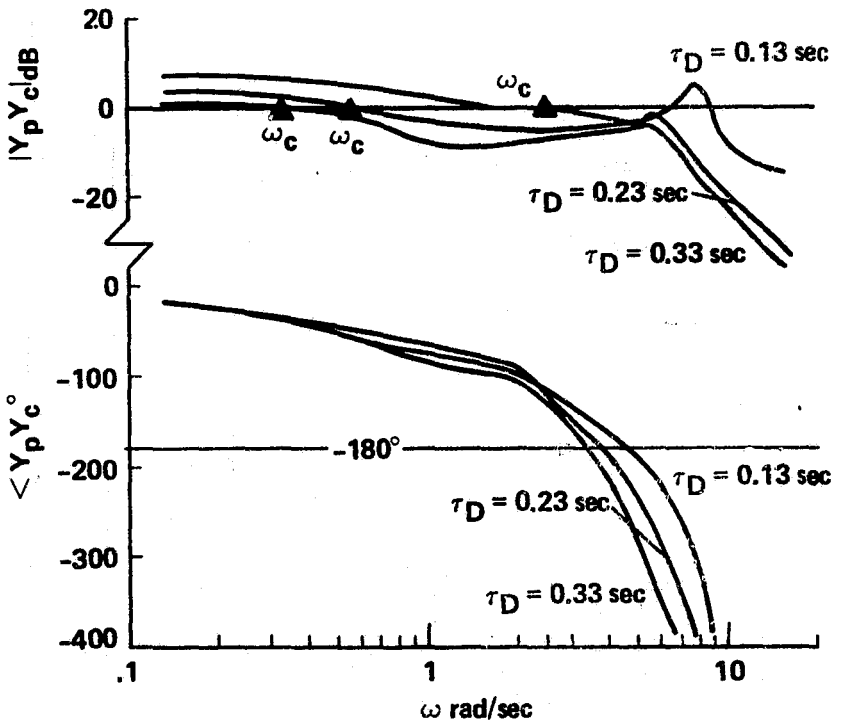


Figure 3. Pilot/Vehicle Dynamics for Three Configurations From Ref. 2.

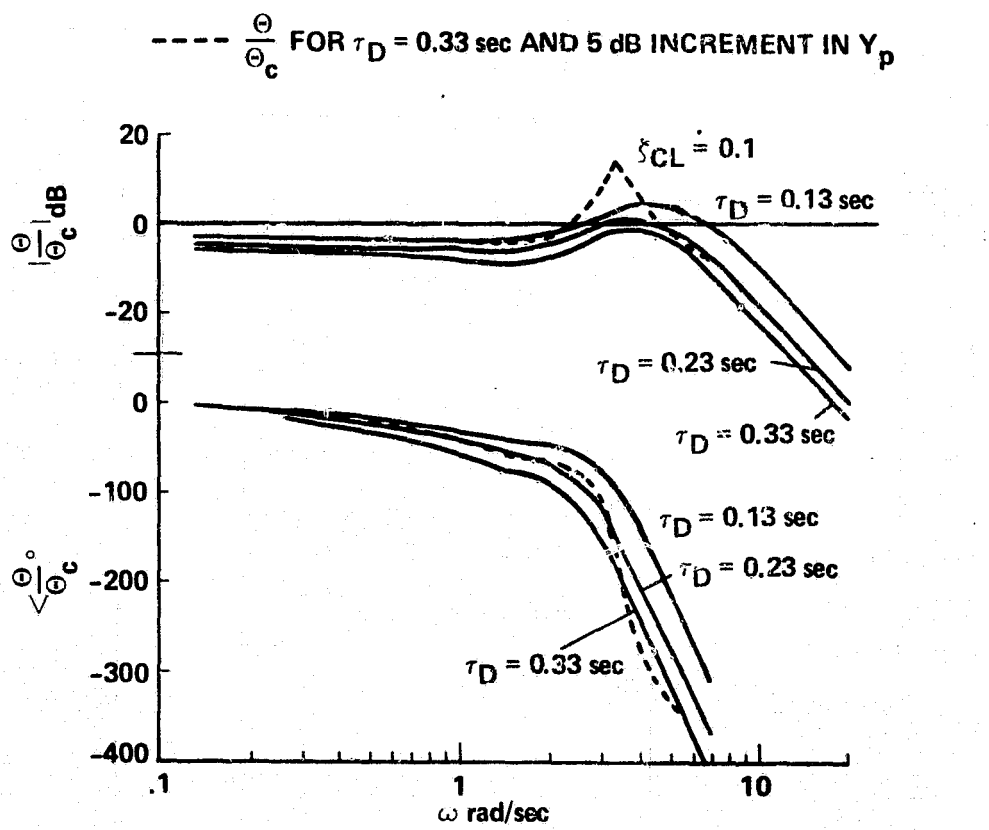


Figure 4. Closed-Loop Characteristics for Three Configurations From Ref. 2.

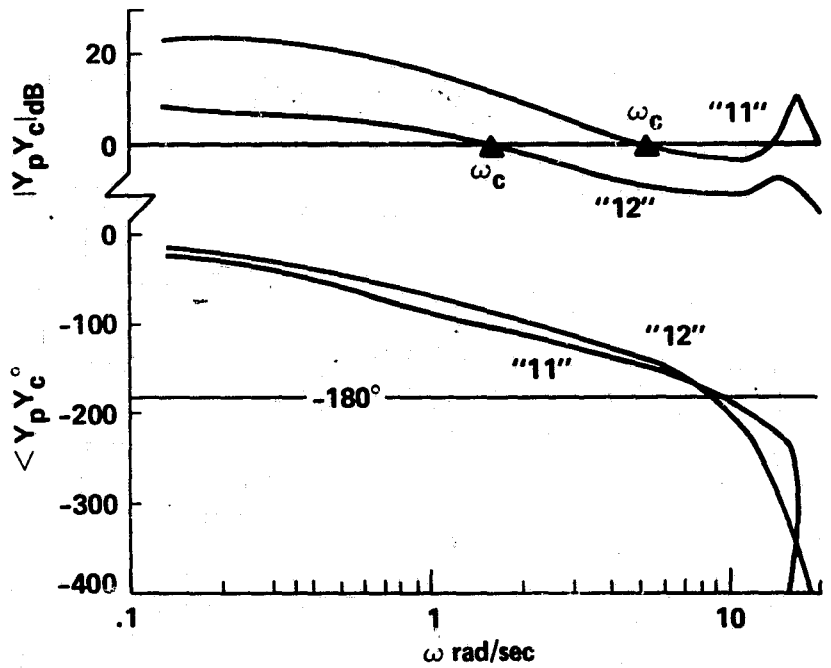


Figure 5. Pilot/Vehicle Dynamics for Two Configurations From Ref. 5.

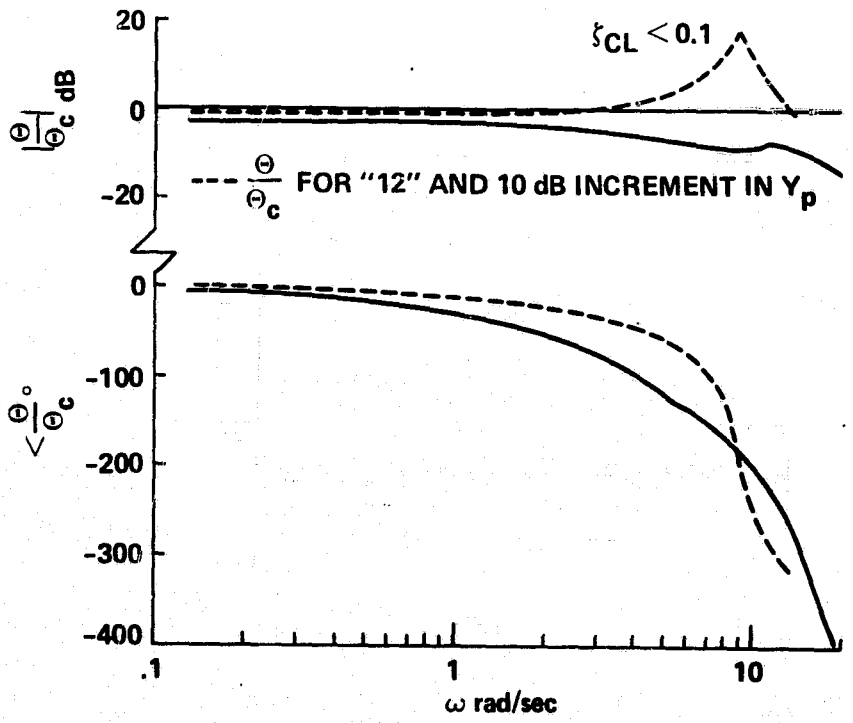


Figure 6. Closed-Loop Characteristics for a Configuration From Ref. 5.

C-4

PIO TENDENCY RATING SCALE

DESCRIPTION	NUMERICAL RATING
NO TENDENCY FOR PILOT TO INDUCE UNDESIRABLE MOTIONS.	1
UNDESIRABLE MOTIONS TEND TO OCCUR WHEN PILOT INITIATES ABRUPT MANEUVERS OR ATTEMPTS TIGHT CONTROL. THESE MOTIONS CAN BE PREVENTED OR ELIMINATED BY PILOT TECHNIQUE.	2
UNDESIRABLE MOTIONS EASILY INDUCED WHEN PILOT INITIATES ABRUPT MANEUVERS OR ATTEMPTS TIGHT CONTROL. THESE MOTIONS CAN BE PREVENTED OR ELIMINATED BUT ONLY AT SACRIFICE TO TASK PERFORMANCE OR THROUGH CONSIDERABLE PILOT ATTENTION AND EFFORT.	3
OSCILLATIONS TEND TO DEVELOP WHEN PILOT INITIATES ABRUPT MANEUVERS OR ATTEMPTS TIGHT CONTROL. PILOT MUST REDUCE GAIN OR ABANDON TASK TO RECOVER.	4
DIVERGENT OSCILLATIONS TEND TO DEVELOP WHEN PILOT INITIATES ABRUPT MANEUVERS OR ATTEMPTS TIGHT CONTROL. PILOT MUST OPEN LOOP BY RELEASING OR FREEZING THE STICK.	5
DISTURBANCE OR NORMAL PILOT CONTROL MAY CAUSE DIVERGENT OSCILLATION. PILOT MUST OPEN CONTROL LOOP BY RELEASING OR FREEZING THE STICK.	6

Figure 7. The Pilot-Induced-Oscillation Rating Scale.

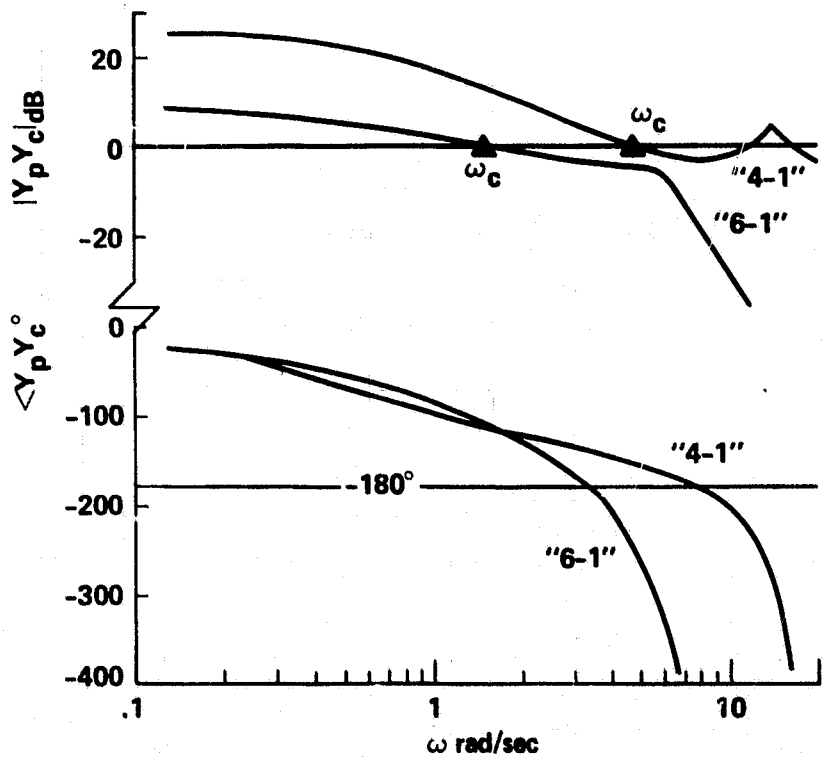


Figure 8. Pilot/Vehicle Dynamics for Two Configurations From Ref. 6.

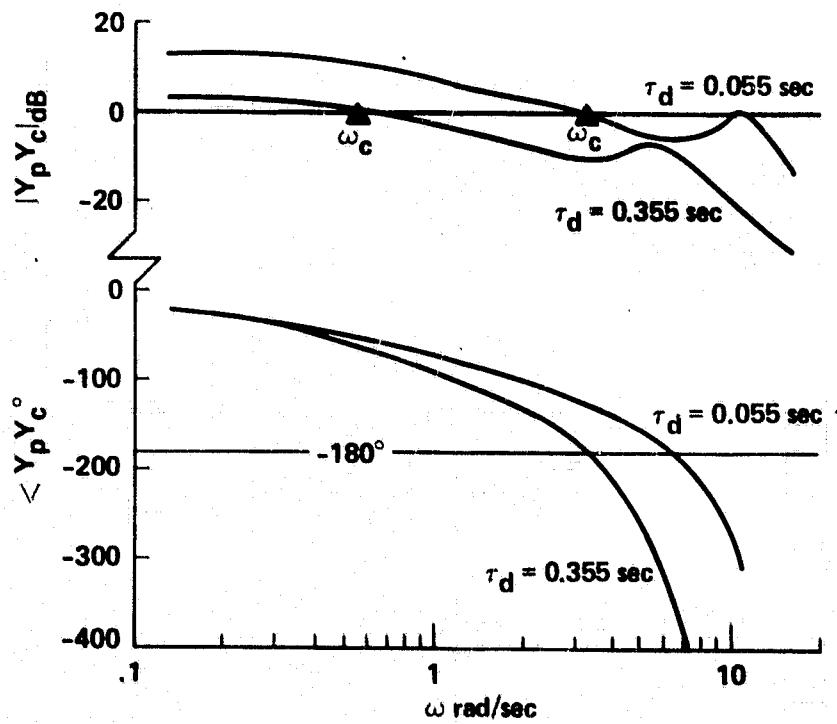


Figure 9. Pilot/Vehicle Dynamics for Two Configurations From Ref. 9.

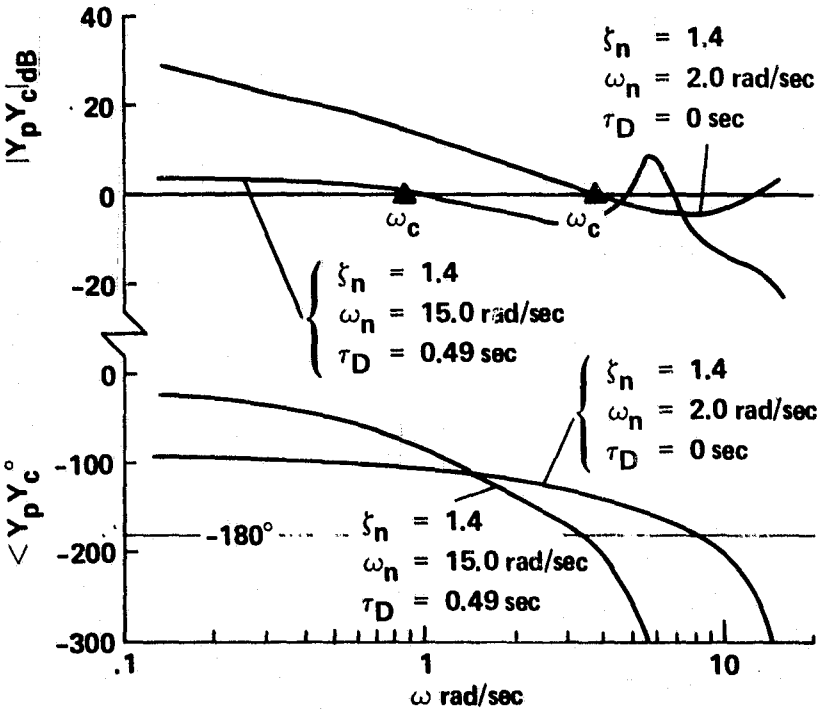


Figure 10. Pilot/Vehicle Dynamics for Two Configurations From Ref. 10.

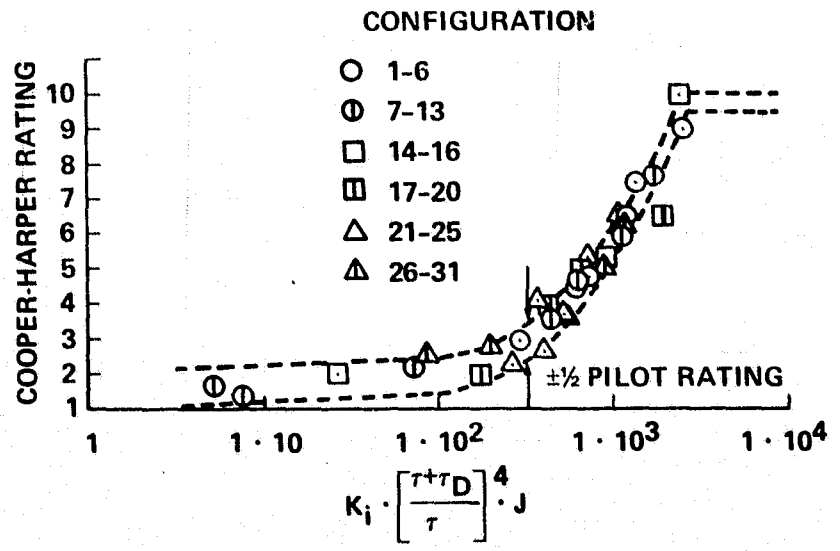


Figure 11. Cooper Harper Pilot Ratings vs a Proposed Model-Based Metric.

D24

A MODEL FOR THE SUBMARINE DEPTHKEEPING TEAM

John R. Ware, John F. Best, Pamela J. Bozzi

ORI, Inc.

and

David W. Kleinman

University of Connecticut

SUMMARY

One of the few real examples that exist of cooperative control of a relatively fast (compared to economic systems) system occurs in the depth control loop of a submarine. Two operators sit side by side viewing essentially similar displays in an arrangement that is much like a conventional aircraft cockpit. One of these operators has direct control of the forward planes and the other has direct control of the stern planes. By tradition each is assigned a distinct control task; the forward planesman controls depth and the stern planesman controls pitch (attitude). Obviously these controls are not independent and a high degree of coupling exists.

The most difficult task the depthkeeping team must face occurs during periscope-depth operations during which they may be required to maintain a submarine several hundred feet long within a foot of ordered depth and within one-half degree of ordered ptich. The difficulty is compounded by the facts that wave generated forces are extremely high, depth and pitch signals are very "noisy" and submarine speed is such that overall dynamics are slow.

In late 1979 we began a study leading to a mathematical simulation of the depthkeeping team based on the optimal control models that have proved successful in many other applications. This work will be described, including a solution of the optimal team control problem with an output control restriction (limited display to each controller).

I. INTRODUCTION

Modern submarines bear little resemblance to their WW I and II ancestors. Prior to the advent of the nuclear age, submarines were basically enclosed surface ships capable of submergence for short periods of time and substantially faster on the surface than below. A typical modern submarine is built to operate and function totally submerged for essentially indefinite periods. Circular cross section hulls and improved hydrodynamic shaping has resulted in a reverse of the WW II standards and the new subs are much faster below the surface than while surfaced.

Despite substantial advances in almost every other area of submarine operation, steering and diving control procedures have changed relatively very little over the years. True the displays are high reliability devices and the planes/rudder hydraulic systems are faster and quieter but man functions in the loop much as he did before. In fact, the steering and diving control of a modern U.S. Navy submarine is one of the few examples of a team effort for the control of a small dynamic system. In this paper we will discuss the most demanding of the functions of the steering and diving team; that is, accurate maintenance of depth, trim (pitch), and heading while operating at periscope depth in a heavy sea way; and our efforts to model their procedures. The remainder of this paper contains the following material:

Section II describes the Near-Surface Depthkeeping (NSDK) control problem in more detail and discusses the roles of the team members.

Section III describes an Optimal Control Modeling approach for solution of this problem.

Section IV presents the results of a preliminary experimental study and describes the follow-on efforts.

II. THE SUBMARINE CONTROL PROBLEM

Over the past several years we have been investigating many aspects of submarine control including display effects and workload. However, the most difficult task, from the operators' viewpoint, is the periscope depth operation. There are several sources of difficulty including:

(1) Control of a large physical system (perhaps 300 to 400 feet) with very slow dynamics.

(2) A narrow operational band; too deep and observation capability is lost, too shallow and detection may result.

(3) The presence of extremely large disturbance forces due to sea waves. These are of two types: first order forces which resemble zero-mean, narrow band, Gaussian processes; and second order forces which resemble the first order

processes past through a square law detector (squaring device). The second order forces always act upwards and tend to "suck" the submarine towards the surface.

(4) The presence of (relatively) high frequency noise on all displays which is the direct or indirect result of first order wave action.

(5) A relatively low skill level and rating (compared to pilots of military aircraft) of the helmsman/planesmen.

In order to cope with these difficulties the U.S. Navy has evolved a procedure that has as its goal the de-coupling of the primary control axes of depth, pitch and heading. There are five personnel involved in the near-surface depthkeeping (NSDK) operation whose duties and responsibilities are as follows:

Officer of the Deck: The OOD is generally responsible for the overall well being of the ship. He will issue orders directly to the Diving Officer of the Watch for depth and trim. He also issues orders directly to the helmsman for heading control. A typical set of orders would be: "Helmsman, xx degrees right rudder, steady on course xxx degrees." The helmsman replies: "Helm aye" and then, when the course is reached, "Steady on course xxx degrees."

Diving Officer of the Watch: The job of the DOOW is to attain and maintain ordered depth. He does this be exerting direct control over three other personnel--the stern planesman, the fairwater planesman, and the Chief of the Watch. One of his major roles is training and admonition.

Chief of the Watch: The COW is in charge of the Ballast Control Panel (BCP) and he ballasts and trims the ship in response to orders from the DOOW. He does this moving water fore and aft (for trim) and to and from the sea (for ballast).

Stern Planesman: The role of the stern planesman is to maintain ship's angle (pitch) using the stern planes (located near the rudder) in direct response to commands from the DOOW.

Fairwater Planesman: Under normal conditions the fairwater * planesman has two separate and distinct responsibilities.

---

* The bridge fairwater, now called the "fairwater" or "sail" is what was called the "conning tower" on the older ships. The fairwater planes are mounted on this structure so that they do not extend beyond the beam of the ship.

First, in his role as planesman, he will be tasked to maintain ship's depth in direct response to orders from the DOOW. Second, in his role as helmsman, he is responsible for attaining and maintaining ship's heading in direct response to orders from the COD.

The physical arrangement of the personnel is as follows. the OOD moves freely through the control room but is usually fairly near the periscope (which is located amidships) from which position he can observe all of the control activity. Slightly forward and to port (left) is the diving control station which somewhat resembles the pilot/co-pilot seats, controls, and displays in an aircraft (no windows, of course). The DOOW sits slightly behind and centered on the planesmen. The Ballast Control Panel is still further to port and the COW sits facing this panel which can also be easily monitored by the DOOW.

The fairwater planesman's role is obviously the more difficult of the two as he receives orders from two sources to control two axes of motion. Fortunately the vertical and horizontal plane motions are only very lightly coupled and, for all practical purposes, he may consider them to be orthogonal. However, the pitch and depth motions are very closely coupled via the dynamics of the ship and correction of depth errors by the fairwater planesman influence pitch angle and the correction of pitch errors by the stern planesman influence depth. The nature of this interaction and the manner in which the planesman solve their problem was the primary interest of our study.

In order to focus the intensity of the effort it was decided to attempt to model only steady state operation and exclude the evolutions involved in coming to the surface and submerging after the operations were complete. As a result of extensive interviews and at-sea observation of procedures we concluded that, again emphasizing steady state operational conditions, the OOD and the COW contributed little to the moment-to-moment control activity. The OOD, once having given orders for depth and heading, would not interfere in the routine unless unusual circumstances occurred. The COW, having trimmed and ballasted the ship to the DOOW's satisfaction, would only monitor ship's status. Therefore our preliminary experiments included only the DOOW and the two planesmen.

From further observation of crews during the preliminary experiments, augmented by further interview data, we concluded that the role of the DOOW was primarily admonitory. That is, he intensified the efforts of the planemen by issuing commands/warnings such as:

* Watch your depth!
* It's starting to come up!
* Maintain your angle!

and so forth. From this we concluded that, were the planesmen more skilled, the role of the DOOW during steady state operation would be minimal. This was confirmed by noting that there was considerably

less communication in low sea states than in high sea states and that little or no communication occurred between the DOOW and the planesmen when the planesmen had considerable experience. The DOOW transferred his much more extensive skill and experience to the planesmen and also served to motivate them should their attention lag. Therefore, we reasoned, the model could take into account the effects of the DOOW by simply employing a slightly higher "skill factor" or attention parameter than would otherwise have been used. The DOOW was also eliminated from the modeling procedure.

Having reduced the diving team to its two essential members it only remained for us to quantify the nature of the inter-relationships between them. We again intended to rely heavily upon interview data to provide insights in this regard. Indeed all of our interviewees (most of whom had, at one time or another, actually been involved in planesman/helmsman training) emphasized the importance of cooperation between the planesmen. Both during training and at-sea operation, planesmen are urged and encouraged to "cooperate" and not to "fight" each other. However, when we attempted to pin down exactly in what way the planesmen were to cooperate we received no satisfactory answers. Apparently "cooperation" was a vague concept and the over-riding rule was an attempt at independent control of depth and pitch. We decided to make the tentative assumption, until the data should prove otherwise, that cooperation, as such, was a myth. That is, while each planesman is aware of the other's activities, he does not attempt to aid the other in achieving his performance goals. Each planesman establishes his own performance goals and attempts to meet these while in control of a dynamic system which is made up of both the ship and the other planesman. In fact, we postulated that should the planesmen be completely isolated from each other and from displays of the other's behavior and controlled states, their responses would not be substantially different. Unfortunately we did not have the opportunity to design and conduct a set of experiments which would have conclusively demonstrated the truth of falsity of that assumption.

III. FORMULATION OF THE MODEL STRUCTURE

The conclusion that cooperation is largely a myth heavily influences the resulting model structure. Beginning basically with the Optimal Control Model (OCM) as proposed by Kleinman (Ref. 1), the model structure shown in Figure 1 was derived. Having had good success with an OCM approach for single person submarine control in the past we elected to continue along this path as far as practicable. For example, the lack of cooperation implies that the planesmen have different performance functionals (control goals). In addition there are other items to be considered with regard to observation and utilization of data. The most basic of these are summarized in Table 1. In Table 2 we present further differences between the human and the optimal controller. As these latter limitations and differences are well known they require no further elaboration at this point.

Figure 1

OPTIMAL CONTROL MODEL (OCM) FOR SUBMARINE NSDK EVALUATION

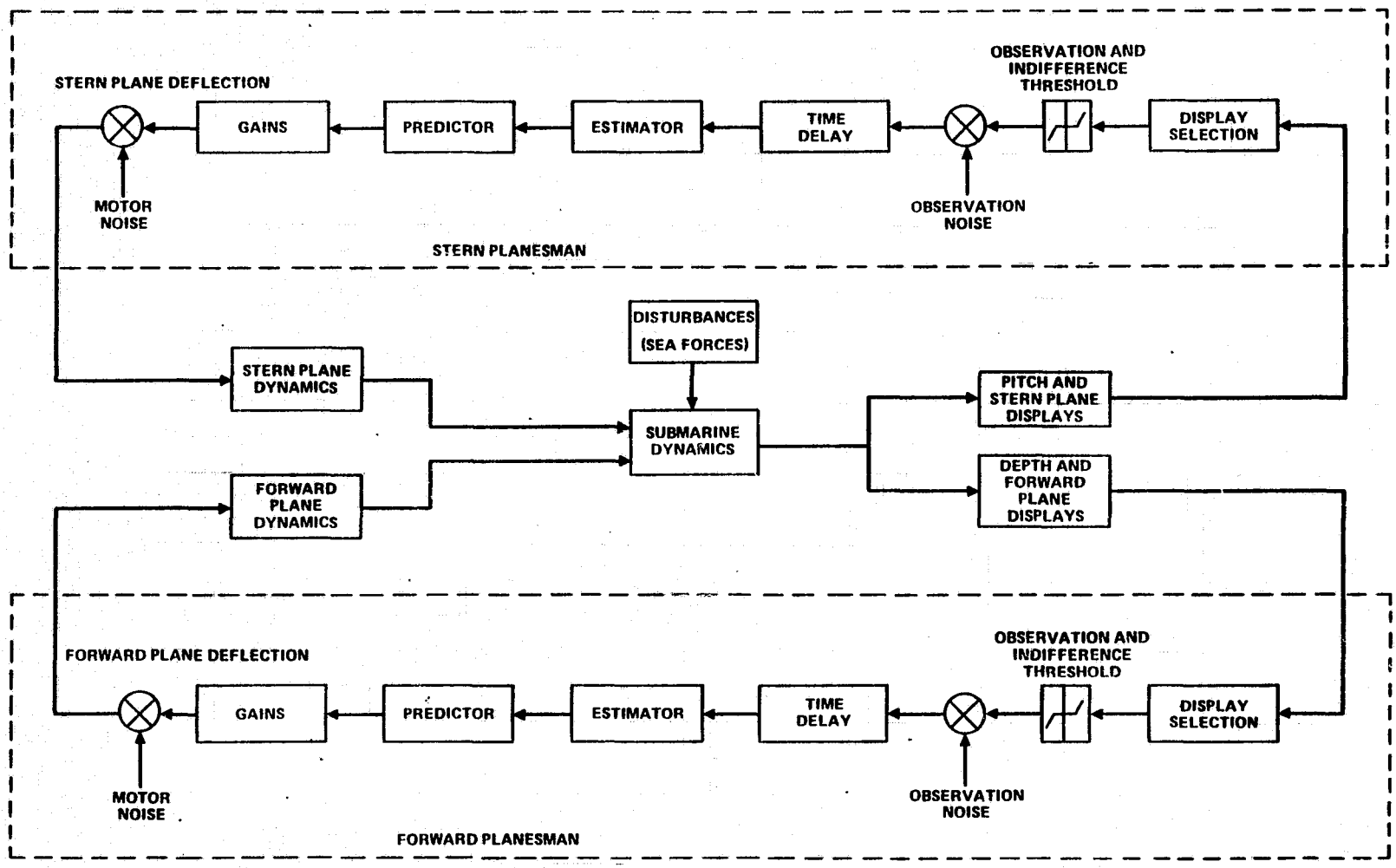


TABLE 1

COMPARISON OF OPTIMAL AND HUMAN TEAM STRATEGIES
DURING NEAR SURFACE DEPTHKEEPING

Optimal Controller Strategy	Human Strategy
The optimal controller considers that both planes operate in harmony and utilizes a single performance function to derive its control law.	The fairwater and stern planesmen's goals (performance functions) are different which may result in conflict.
The optimal controller "observes" all states and uses this information in the control laws.	The fairwater planesman observes depth information only and the stern planesman observes pitch information only. Neither attempts to construct the others states from his limited observations.
The optimal control law uses gains on all states for both controllers.	The forward planesman's control law uses depth and depth rate only while the stern planesman's control law uses pitch and pitch rate only.

TABLE 2

COMPARISON OF OPTIMAL CONTROL AND HUMAN LIMITATIONS

Optimal Controller	Human Constraint
Optimal system's observations are "noise free" (neglecting sensor noise).	Humans observe imperfect instruments and are subject to observation and indifference threshold phenomenon.
Optimal controller has perfect knowledge of submarine dynamics.	Human's knowledge of submarine dynamics is imperfect and a strong function of training and experience.
The optimal controller observes all states simultaneously and is able to utilize the information so obtained.	The human can effectively observe only one display at a time and must allocate his attention appropriately.
The optimal controller gives full "attention" to the task at all times.	Humans are subject to lack of attention due to mental fatigue, boredom, stress, or lack of motivation.

The structure shown in Figure 1 is at variance with the standard Linear-Quadratic (LQ) controller (and the human models resulting from it) in two major respects:

(1) Each control has its own performance function. In the present case these are independent.

(2) The control law is based on a limited set of states.

The combination of these deviations from standard LQ control theory is one aspect of team theory and/or decentralized control theory. While literature on these topics is easily found little or no practical applications or solution techniques is evident. In the limited space here we could not present the details of the techniques we used to solve this particular problem. An outline of the approach may be, however, instructive. An initial set of gains is determined, by an appropriate start-up technique, for one of the controllers. This set of gains, plus the system open loop dynamics, represent the system to be controlled by the other controller whose gains are then easily computed by standard techniques. This set of gains is then used to update the original guess on the first controller's gains and the process is repeated iteratively until sufficient convergence takes place. This iterative procedure is embedded in a further iterative procedure (due to Wenk, Ref. 2) for computing the optimal output control (state limited) feedback problem.

Having determined a computational procedure for solving the requisite control problem there remained only the selection of the various model parameters. These are:

(1) The control and state weighting matrices.
(2) The indifference thresholds.
(3) The Total Attention Parameter (TAP).

The Total Attention Parameter (TAP) is simply a divisor of the baseline observation noise derived from previous display related experiments (-20.0 db). By varying the TAP from unity towards zero the observation noise increases which, in turn, decreases the performance of the model. This decrease in performance may be thought of as arising from any of several causes such as boredom, lack of motivation, and so forth. Based on previous experiments we expected to be able to match the human data with a TAP in the range of 0.5 to 0.75. The diagonal elements of the weighting matrices would simply be the inverse of the variances of the measured human performance and thresholds would be approximately one-half of display graduations.

Unfortunately this straight-forward procedure did not duplicate the experimental data available. In every case the performance of the model was substantially superior to crew performance unless the TAP was reduced to unreasonable levels. After a considerable period of investigation it was concluded that the error did not reside with our assumed model structure and gain selection procedures. However, based to a large extent on detailed studies of strip chart records, it was

postulated that the lower than expected performance of the crews was attributable to a relative lack of motivation which was manifested by higher than normal indifference thresholds. Because it is not unreasonable to suppose that the TAP and the indifference thresholds would be related when motivation was low, we sought to form an ad hoc relationship between the two that would explain the available data. The form of this relationship was suggested by the form of the Random Input Describing Function (RIDF) for the threshold in the range of interest. The TAP and the RIDF for the threshold interact in such a way as to imply that they could be related by:

$$\text{Indifference Threshold}_i = K_{1i} - K_{2i}(\text{TAP})^2$$

It should be amphasized that this relationship was selected for mathematical tractability and may have to true anthropomorphic basis.

The K_{ni} were selected by (the admittedly) arbitrary procedure of assuming that the normal threshold (½ major graduations) could be used for a TAP of 0.8 and that the thresholds would be four times as great when the TAP was 0.2. It is only necessary to vary the TAP until the desired level of performance is attained. The validity of the approach could only be determined by its ability to predict performance in conditions other than that used for tuning.

IV. PRELIMINARY EXPERIMENTAL RESULTS

A preliminary experiment was conducted for the purpose of providing data for model tuning. In this preliminary experiment only the vertical plane states of depth and pitch were controlled. (The total control problem will be addressed in a later study). Eight crews controlled a simulated U.S. Navy attack class submarine under two conditions; Sea State A at X knots and Set State B at Y knots. For the first case the TAP of the model was varied until the model's average RMS depth error equalled that of the average RMS depth error of all crews. Weighting matrices based on this condition were used to generate the data shown in Table 3A. (All data has been normalized so that the value for the human results is unity for every variable in order to preserve the unclassified status of this report.) By agreement with the Navy we are only required to match pitch and depth data but we have reported control data also. Exactly the same model parameters were used to predict performance in the other condition, Sea State B at Y knots, as shown in Table 3B.

V. CONCLUSION

Based on the preliminary results it certainly appears that our decision to proceed using an OCM approach was more than justified. The team modeling problem can be very complex but the OCM structure allowed us to attack the problem in a fairly coherent manner. At the present time the Navy is planning a much more elaborate set of experiments which

TABLE 3

COMPARISON OF MODEL AND HUMAN PERFORMANCE

TABLE 3A

SEA STATE A AT X KNOTS
(HUMAN DATA IN PARENTHESES)

	Depth Error	Pitch Error	Stern Plane	Fairwater Plane
Average RMS	0.990 (1.000)	1.007 (1.000)	1.149 (1.000)	1.019 (1.000)
Standard Deviation	0.076 (0.073)	0.065 (0.073)	0.072 (0.362)	0.070 (0.161)

TABLE 3B

SEA STATE B AT Y KNOTS

	Depth Error	Pitch Error	Stern Plane	Fairwater Plane
Average RMS	0.991 (1.000)	0.929 (1.000)	0.679 (1.000)	0.684 (1.000)
Standard Deviation	0.123 (0.097)	0.079 (0.106)	0.057 (0.273)	0.041 (0.188)

will include both depth and heading control along with other variables to fully validate the OCM approach. Perhaps we'll be back to tell about this next year.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support and encouragement of Mr. Gary Jones and Mr. William Louis of the Naval Sea Systems Command. Mr. Leo Hayes and Dr. Ronald Offenstein of the Marine Systems Division of Autonetics also provided many useful suggestions and designed, conducted, and provided data for the preliminary experiment.

REFERENCES

1. Kleinman, D.L., et. at., "A Control Theoretic Approach to Manned-Vehicle Systems Analysis," IEEE Transactions on AC, V-1, AC-16, No. 6, 1971.

2. Wenk, C.J., "Parameter Optimization for Linear Systems with Arbitrarily Constrained Information and Control Structure", University of Connecticut School of Engineering Technical Report TR 79-4, February 1979.

MODELING OF THE AIRCRAFT IN-TRAIL-FOLLOWING TASK DURING PROFILE DESCENT*

T. Goka, J.A. Sorensen and A.V. Phatak

Analytical Mechanics Associates, Inc.
2483 Old Middlefield Way
Mountain View, CA

ABSTRACT

The FAA and NASA are jointly developing Cockpit Display of Traffic Information (CDTI) system concepts which enable the pilot to observe the surrounding air traffic pattern. The impact of such a system is far reaching in terms of improved safety, pilot and controller workload, and aircraft fuel efficiency. One direct payoff is the ability to distribute the ATC workload to the pilot in such tasks as merging and spacing. In this paper, the CDTI application of spacing approach aircraft in the terminal area is addressed. At both Langley and Ames Research Centers, in-trail-following/CDTI experiments were performed using realistic cockpit simulators and profile descent approach scenarios. Based on collected experimental simulator data, pilot models were developed which include state estimation, decision making and flight control aspects. These models were coupled with models of aircraft and CDTI equipment to study the dynamic phenomena and stability of strings of aircraft along various approach patterns.

INTRODUCTION

Both the use of more automation and more involvement of the pilot in the air traffic control process are well understood to be future needs for providing greater terminal area capacity. In hearings conducted in June 1977 by the U.S. House of Representatives Subcommittee on Transportation, Aviation and Weather it was concluded [1] that NASA should make extensive use of its Terminal Configured Vehicle (TCV) and other cockpit simulators to assist the FAA in:

(1) examining the capabilities and limitations of cockpit displays of traffic information (CDTI);

(2) exploring distributive management concepts for air traffic control; and

(3) examining human factor problems related to distributive management concepts.

Since that time, a program has been organized which will tie together FAA ATC (ground-based) simulators and NASA aircraft and associated cockpit simulators into a joint research project [2] to explore applications of the CDTI system.

One application of particular interest is the use of the CDTI display by the pilot for non-vectored clearances relative to other traffic. Under this category are functions such as control into a traffic merge point, and queuing, or spacing, along a route.

---

* The research described in this paper was supported by NASA Langley Research Center under Contract No. NAS1-16135.

In order to derive the control requirements for such functions, it is first necessary to understand the dynamics of merging and trailing aircraft.

Several questions arise associated with this CDTI-based terminal area traffic tactical control concept. These include:

(1) What are the basic dynamic phenomena associated with independently controlled strings of aircraft?

(2) What conditions would produce instability in the string?

(3) What information does each pilot need (from the CDTI and elsewhere) to merge his aircraft adequately into the string and then to maintain appropriate spacing?

(4) What are the effects of measurement and display errors, wind shears, aircraft mixes, spacing constraints, and merge trajectories on the dynamics and control performance of the system?

(5) What advantages does this concept have compared to the ground-based control?

This study begins to address these questions from a systems point of view.

In this first year's effort, focus was placed on analysis of the dynamics of already formed strings of aircraft. To aid in this analysis, use was made of data from NASA cockpit simulator studies of in-trail following. The experimental data were used to confirm analytical predictions and to uncover new phenomena for the spacing task.

BACKGROUND

System Overview

The flight system (i.e. the pilot/aircraft/CDTI combination interacting with other aircraft and ATC) is assumed to be entering the terminal area and proceeding along an established approach to landing. A sketch of such a scenario is depicted in Fig. 1. The flight systems can be further described by the block

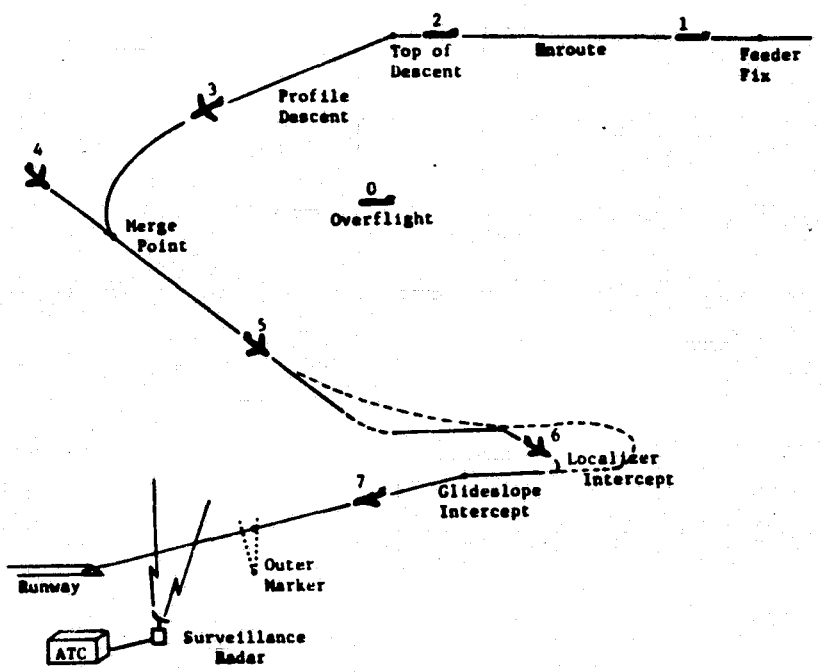


Figure 1. Sketch of Approaching Aircraft in a Terminal Area Scenario.

diagram shown in Fig. 2. With regard to the ten blocks of Fig. 2, the following assumptions were made for this initial effort:

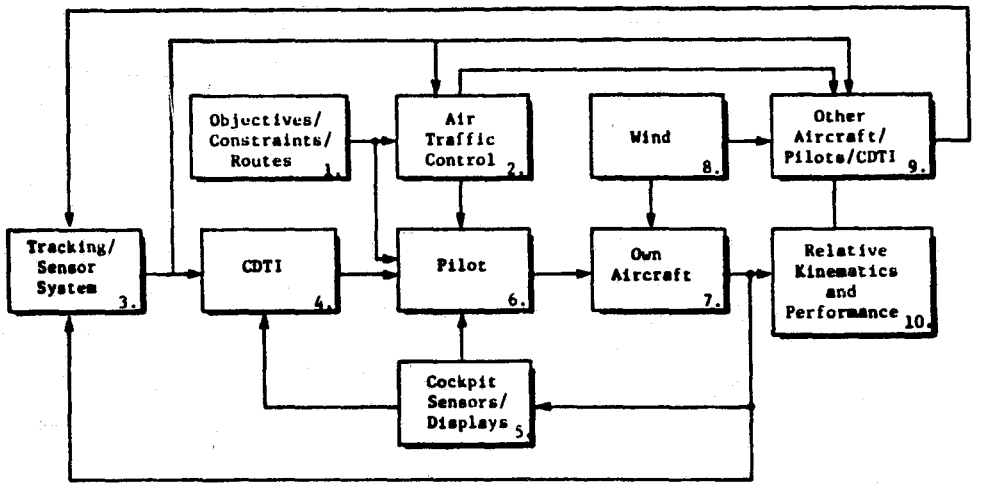


Figure 2. Block Diagram of a Flight System with CDTI Equipment.

1. Lead aircraft were given specified descent profiles for the approach task. Profile descents into Denver Stapleton International Airport were examined.
2. Air Traffic Control was used to issue the initial sequencing and separation commands (e.g., "Close and maintain 60 sec separation behing UA Flight 305").
3. No specific tracking system was assumed. Target aircraft were displayed as symbols which took discrete movements every four sec.
4. The CDTI display symbology was based on formats chosen for the NASA cockpit simulators. [3, 4, 5]
5. Other cockpit displays were those standard for the simulated aircraft.
6. Part of the modeling objective was to characterize the pilot using the CDTI in a following situation to sufficient detail so that the overall string dynamics could be accurately emulated. Thus, only the longitudinal control was particularly studied.
7. Modeling of own aircraft longitudinal dynamics was relatively simple.
8. No wind model was used.
9. Other aircraft trajectories were stored from previous simulations.
10. The relative longitudinal dynamics of each aircraft was measured as projected on the desired horizontal flight path.

Previous Vehicle String Studies

There has been considerable previous investigation of traffic flow and control problems of ground vehicles in strings [6-8]. Probably L.A. Pipes' work of the early 1950's [6] was the first attempt utilizing the methods of operations research. He derived a mathematical model for strings of automobiles (which was a basic model used by later researchers), and he studied dynamic behavior of a string of vehicles initially at rest or after a sudden stop of the leading

vehicle. Haight [5], contributed a great deal to the understanding of traffic flow by assuming a stochastic environment and using queuing theory. However, because his approach was macroscopic in nature, such problems as to how to control individual vehicles or how the stability of the string of vehicles is affected were not resolved.

Athans and others [9-10] solved the optimal control problem of a string of vehicles via the well known LQG (linear, quadratic and Gaussian) method. Athans and Porter [10] applied these techniques to the problem of controlling aircraft in the near terminal area under somewhat restrictive assumptions. The LQG approach is mathematically interesting and concise; however, it is very difficult to realize in light of CDTI applications.

Another approach to the metering and spacing problem was that taken by Tobias [11]. He obtained a general scheduling algorithm to generate time slots at each way point for each aircraft traveling along the air route. However, because his simulation did not provide dynamic interactions between adjacent aircraft, it cannot be utilized to study string stability.

Based on a review of the above work and other pilot modeling efforts [12-14], it was determined that a fresh start was needed to understand the dynamic phenomena and stability aspects of a string of decelerating, descending aircraft in a terminal area. This required analyses of different possible separation criteria and the development of longitudinal flight system models.

SEPARATION CRITERIA

Finding a suitable longitudinal distance separation criterion is especially important when the pilot independently executes the spacing task with the aid of a CDTI system. The separation criteria must satisfy three qualifications: safety/efficiency, executability, and computability/displayability. There could be numerous criteria which satisfy these requirements based on either distance, ground speed, or time. Four possible criteria expressed mathematically are:

(1) constant distance (CD)

$$\Delta d = d_i - d_{i+1} = \text{constant} ,$$

(2) constant time follower (CTF);

$$\Delta d = d_i - d_{i+1} = T_F V_i \quad (T_F = \text{constant}) ,$$

(3) constant time predictor (CTP);

$$\Delta d = d_i - d_{i+1} = T_P V_{i+1} \quad (T_P = \text{constant}),$$

(4) constant time delay (CTD);

$$d_i(t) - d_{i+1}(t-T_D) = 0 \quad (T_D = \text{constant}) .$$

Here, Δd is the separation, d_i and d_{i+1} are longitudinal distances of the ith and (i+1)th aircraft from a common reference point, and V_i and V_{i+1} are the corresponding ground speeds.

Differentiating the separation distance criteria yields the ideal speed profile to maintain the separation. This neglects pilot and aircraft caused delays. For example, in the case of the constant distance criterion, the separation equation is given by:

$$d_i(t) - d_{i+1}(t) = \text{constant} \tag{1}$$

Differentiating this equation yields

$$V_{i+1}(t) = V_i(t) \ . \tag{2}$$

Equation (2) implies that in order to maintain a constant separation distance, the following aircraft speed must be identical to that of the lead aircraft. Table 1 summarizes the ideal speed profile corresponding to each of the five separation criteria listed above. The table also shows the effect of each criterion repeated for the ith time (as in a string) with respect to the first aircraft speed.

Table 1. Separation Criteria and Ideal Velocity Profile

Separation Criteria		Ideal Velocity Profile	i^{th} Vehicle
Constant Distance (CD)	$d_i - d_{i+1} =$ constant	$v_{i+1}(t) = v_i(t)$ $v_{i+1}(s) = v_i(s)$*	$v_i(s) = v_1(s)$
Constant Time Follower (CTF)	$d_i - d_{i+1} = T_F v_i$, T_F = constant	$v_{i+1}(t) = v_i(t) - T_F \dot{v}_i(t)$ $v_{i+1}(s) = (1 - T_F s) v_i(s)$	$v_i(s) = (1 - T_F s)^{i-1} v_1(s)$
Constant Time Predictor (CTP)	$d_i - d_{i+1} = T_p v_{i+1}$, T_p = constant	$\dot{v}_{i+1}(t) = -\frac{1}{T_p} v_{i+1}(t) + \frac{1}{T_p} v_i(t)$ $v_{i+1}(s) = \frac{1}{T_p s+1} v_i(s)$	$v_i(s) = \left(\frac{1}{T_p s+1}\right)^{i-1} v_1(s)$
Combination Constant Time Predictor-Follower (CTFP)	$d_i - d_{i+1} = T_F v_i + T_p v_{i+1}$ T_F, T_p = constants	$\dot{v}_{i+1}(t) = -\frac{1}{T_p} v_{i+1} + \frac{1}{T_p} [v_i - T_F \dot{v}_i]$ $v_{i+1}(s) = \frac{1-T_F s}{T_p s+1} v_i(s)$	$v_i(s) = \left(\frac{1-T_F s}{T_p s+1}\right)^{i-1} v_1(s)$
Constant Time Delay (CTD)	$d_{i+1}(t) - d_i(t-T_D) = 0$, T_D = constant	$v_{i+1}(t) = v_i(t-T_D)$ $v_{i+1}(s) = e^{-T_D s} v_i(s)$	$v_i(s) = e^{-(i-1)T_D s} v_1(s)$

* s = Laplace Operator

IN-TRAIL FOLLOWING EXPERIMENT RESULTS

Results from experiments conducted at NASA Langley Research Center, using their TCV cockpit simulator, were analyzed. The nominal approach path followed by the lead aircraft was along the profile descent leading to Denver Stapleton runway 35R depicted in Fig. 3. The trailing aircraft was situated to begin at the KEANN waypoint. The nominal separation criterion for these experiments was

$$\Delta d = \max [V_o T, 3 \text{ nmi.}] \tag{3}$$

where Δd is the nominal separation, Vo is the own aircraft ground speed, and T is a time constant of 60 sec. This criterion is a combination of the constant time predictor (CTP) and the constant distance (CD) of Table 1, with crossover being at own ground speed of 180 kt.

Experiment Design

For Experiment No. 1, the lead aircraft was placed 7 nmi. in front of the trailing aircraft. This represented a positive initial separation error of 1.7 nmi. for own, when own had an initial ground speed of 340 kt. Own pilot's first task was to close the separation error to 5.3 nmi. After this capture phase, the task was to hold the separation according to Eq. (3) and to maintain the aircraft reasonably close to the nominal profile depicted in Fig. 3.

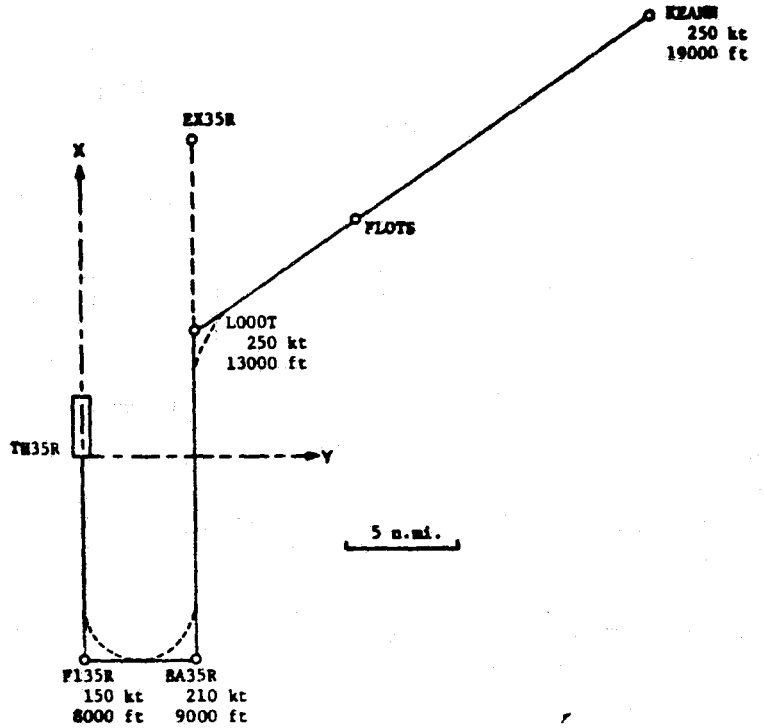


Figure 3. Nominal Approach to Denver Runway 35R

Eight runs, consisting of a single following aircraft trailing nominal leads, were selected for analysis from Experiment No. 1. Longitudinal separation was measured by first projecting the lead aircraft position onto the future position of the following aircraft.

For Experiment No. 2 ("the daisy chain experiment"), eight successive trailing aircraft were initialized at the KEANN waypoint when the immediately preceding aircraft passed 5.3 nmi. (60 sec) ahead of this point. The lead aircraft of this nine-aircraft string followed a nominal profile descent along the path of Fig. 3. There were no initial separation errors in Experiment No. 2. The Eq. (3) separation criterion was again used.

An example of the six key variables that were recorded from each of the experiments are depicted in Fig. 4. These variables are the own and lead aircraft ground speeds, the actual and nominal relative separations between own and the lead aircraft, and own aircraft's throttle and spoiler settings. These data were used to construct two different pilot models characterizing the results of Experiment Nos. 1 and 2.

Pilot Models

In examining Experiment No. 1 data, a great deal of variation in pilot strategy and actions was seen from run to run, even through the same pilot could be following the same lead aircraft with identical initial conditions. Thus, the models developed had to contain features which allowed for changes in pilot decision and control from run to run. The approach taken for developing the combination pilot/aircraft/CDTI flight system model for in-trail following was direct. Given a set of input/output time sequences from the experiments, the model was designed as a functional set of equations and logic with variable parameters which would provide similar input/output sequences.

Fig. 5 depicts a first-level block diagram of the flight system model developed here, which is referred to as Model No. 1. Aircraft control dynamics are represented by a first order lag between the commanded and actual accelerations. The actual acceleration is integrated twice to obtain own aircraft ground speed and distance traveled.

The CDTI display is modeled to generate the separation error $\tilde{\delta}r_M$. The recorded target ground speed V_T is integrated to obtain the target distance traveled, r_T. The distance traveled by the flight system model, r_M, is subtracted from r_T to produce the model separation distance r_{act}. The model ground speed, V_M is multiplied by 60 sec (which is limited to 3 nmi. from below) to compute the model's nominal separation distance r_{Nom}. The model separation r_{act} is subtracted from r_{Nom} to obtain the model separation distance error $\tilde{\delta}r_M$. The pilot sees a quantized value of $\tilde{\delta}r_M$ on the CDTI display. All modeled values are initialized to those

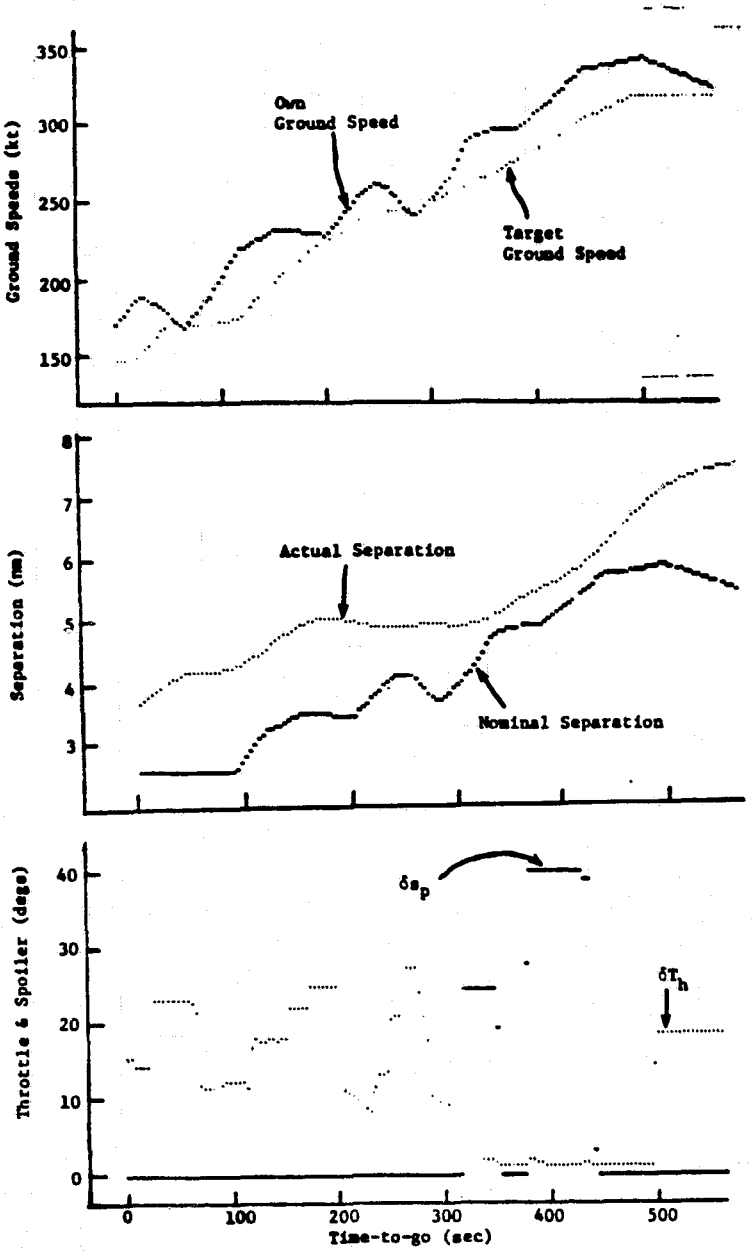


Figure 4. Example of Data Recorded From Experiment and Used to Develop System Model.

recorded during an experimental run. The model separation error is input to a second order tracking filter to approximate the pilot's estimate of the error and its rate ($\Delta\hat{r}$, $\Delta\hat{v}$).

Fig. 6 shows further details of the CDTI/pilot model implementation. A tentative regulator acceleration command is obtained by using a constant gain regulator law based on output from the pilot estimator. That is,

$$a_{Ml} = G_p(\Delta\hat{r} + \Delta r_{Ref}) + G_v\Delta\hat{v}. \quad (4)$$

The regulator law controls the separation error to a bias, Δr_{Ref}. This bias is a function of the map scale and it was introduced based on an observation that pilots tend to stand off at the initial part of the approach and then tend to close in later.

The resulting regulator acceleration is averaged over a 10 sec time interval. If the average value exceeds the current commanded value by a threshold amount θ_a, then the average value $\bar{a}$ is used for the command. This logic simulates the fact that the pilot tends to not change the throttle setting unless the error or command builds up beyond a certain point. Also, he changes the throttle position to a new point which is then held. The acceleration command a_T contains two components: The one, a_{Th} is for the throttle, and the other, a_{sp} is for the spoiler. The acceleration command due to the spoiler provides one discrete deceleration pulse when the separation estimate is closer than the threshold value.

Authority and velocity limiting at the various phases of the approach are introduced to simulate additional observed pilot behavior. For example, the pilot never accelerated above 340 kt nor did he decelerate below 130 kt. regardless of the separation error.

Fig. 7 shows typical time plots of the results of using this flight system model compared to experimental results. The top plot in Fig. 7 compares three ground speeds - recorded target, recorded own, and model predicted own (V_M in Fig. 5). The second plot compares three separation distances between own and target - recorded, model derived (r_{act} in Fig. 5), and nominal model (r_{Nom} in Fig. 5). To produce these results, the model depicted in Figs. 5 and 6 was given identical initial conditions (separation, ground speed) as that of the actual run, and it was driven by the recorded target ground speed V_T.

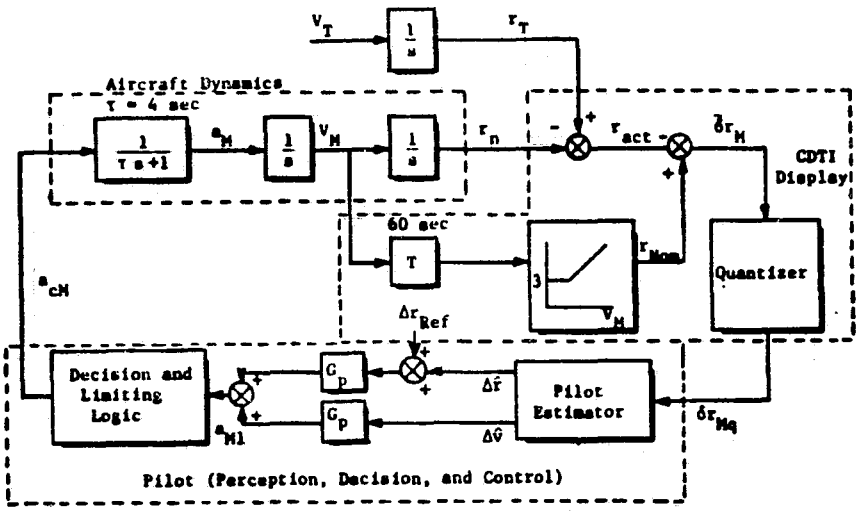


Figure 5.

First Level Block Diagram of In-Trail Following Flight System Model.

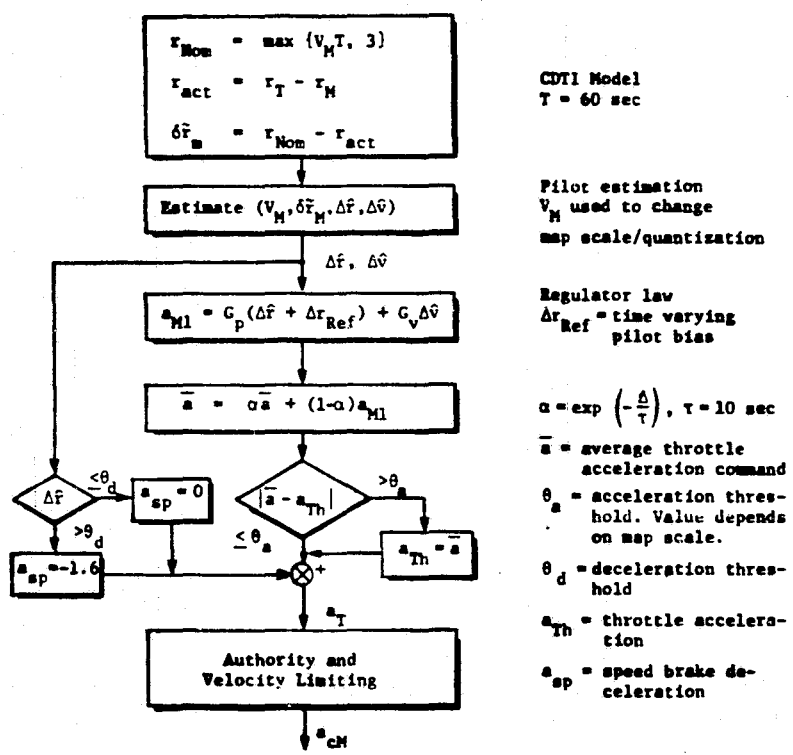


Figure 6.

Second Level Block Diagram of CDTI and Pilot Acceleration Command Generation.

The goodness by which the flight system model matched the performance of the actual experiments varied somewhat from run to run. Over the eight runs, the model predicted ground speed had a mean error of -5.1 kt and an rms error of ±15.3 kt when compared to the recorded ground speed of the following aircraft. Part of the error is due to the fact that the recorded ground speed excursions were greater than predicted by the model. The frequency spectrum of the excursions in the model and actual data were seen to agree well. This indicates that control gain and acceleration command thresholds are good representations of the pilots' control tactics.

In analyzing the Experiment No. 2 data, some observations were made that demonstrated a different following behavior than in Experiment No. 1. To obtain these differences, the flight system Model No. 1 was modified in four ways. Observations and model modifications were as follows:

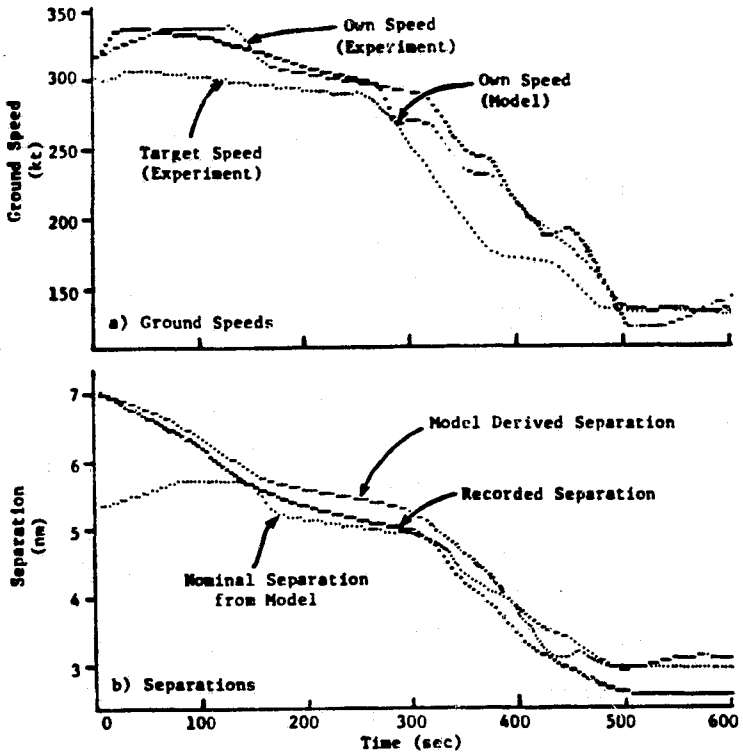


Figure 7.

Comparison of Actual and Model Derived Speeds and Separations.

(1) The initial deceleration of own aircraft was limited to a smaller value than in Experiment No. 1
(2) Own aircraft had a different ground speed undershoot limit when it was decelerating than the 10 kt used for Model No. 1.
(3) The model separation bias Δr_{Ref} was changed to different levels during a run depending on own's ground speed.
(4) The acceleration threshold θ_a shown in Fig. 6 was made smaller to reflect that the pilot's acceleration commands had smaller changes more often. This reflects better tracking accuracy and smaller separation excursions experienced in Experiment No. 2.

The result of these changes is referred to as Model No. 2. The difference between recorded and Model No. 2 ground speeds had a mean error of -0.5 kt and an rms value of ±10.6 kt which is a 30% improvement over the fit provided by Model No. 1.

With a large enough data base, the parameters that are contained within the two models could be treated as stochastic variables; they could then be picked randomly from run to run while exercising the model. However, it is emphasized that our purpose here is not to identify the perfect model but rather to capture the essence of the performance of the pilot/aircraft/CDTI combination in the in-trail following task.

String Dynamic Simulations

The flight system Model Nos. 1 and 2 were used to simulate the longitudinal dynamics of a string of nine aircraft. The lead aircraft in this simulated string followed a sequence of constant decelerations at discrete time points to produce a profile similar to the nominal approach. Initial spacing errors were small.

Fig. 8 shows the simulated results for the first, fifth, and ninth aircraft using Model No. 2. The first plot depicts the following aircraft spacing error, the second plot compares the ground speeds as functions of range-to-go. The interesting point to note from these plots is that each successive aircraft slows

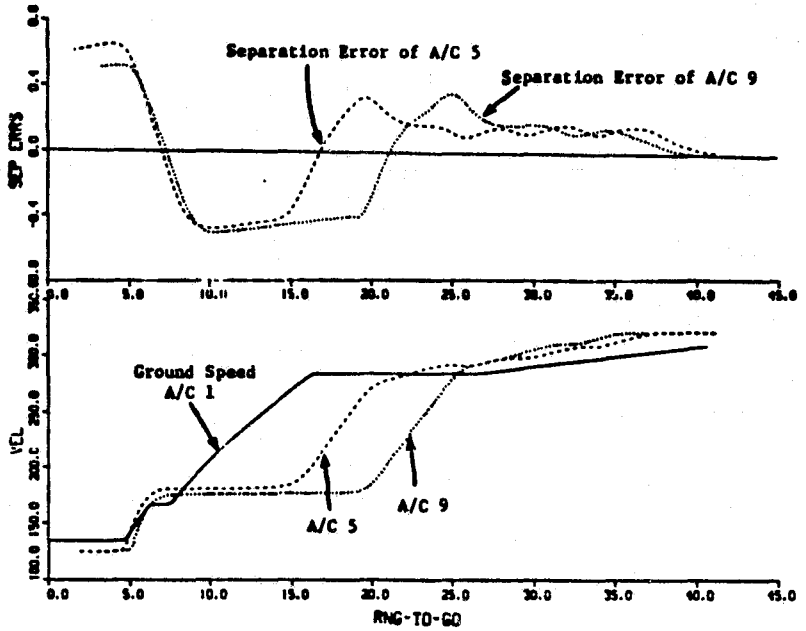


Figure 8.

Comparison of Profiles for Aircraft Nos. 1,5, and 9. Model No. 2.

down at an earlier range-to-go. The slow-down effect was seen on both Experiment No. 2 results and the two simulated strings of nine aircraft using Model Nos. 1 and 2.

The slow-down effect was not caused by the 60 sec time lag inherent in the CTP separation criterion. To prove this conclusion, a simple autopilot was designed (to replace the pilot model) to null out the separation and speed errors. The nominal separation distance was set to the constant time predictor criterion of 60 V_o (i.e., the criterion did not switch at 180 kt.).

The autopilot model was also used to simulate a string of nine aircraft as before. The results of this simulation were then compared to those of the string simulation using Model Nos. 1 and 2 and the actual results from Experiment No. 2. Fig. 9 compares results for the ninth aircraft in each of these four cases. The top plot shows the separation errors, and the bottom plot compares the predicted and actual ground speeds.

Conclusions

From the previous results, the following conclusions can be made:

(1) Both Model Nos. 1 and 2 Experiment No. 2 aircraft have slower ground speeds than the ideal autopilot model, and so they take longer to arrive at the outer marker. (This amounted to an increase in flight time of about 13%.) Thus, the CTP criterion is not responsible for the slow-down.

(2) The differences in separation errors and ground speeds predicted by the Model Nos. 1 and 2 and the ideal autopilot model indicate the addition of in-trail errors caused by the pilots' decisions and actions. The pilot introduces errors because of many items: (a) the switch in the separation criterion and the tendency to hold a greater than indicated separation, (b) inattention to tracking caused by the competitive task of steering to the nominal profile and other tasks associated with landing, and (c) the hunting nature observed in the pilots' acceleration inputs. This latter fact is probably due to the pilot not having a good ground speed reference to use to null the separation error.

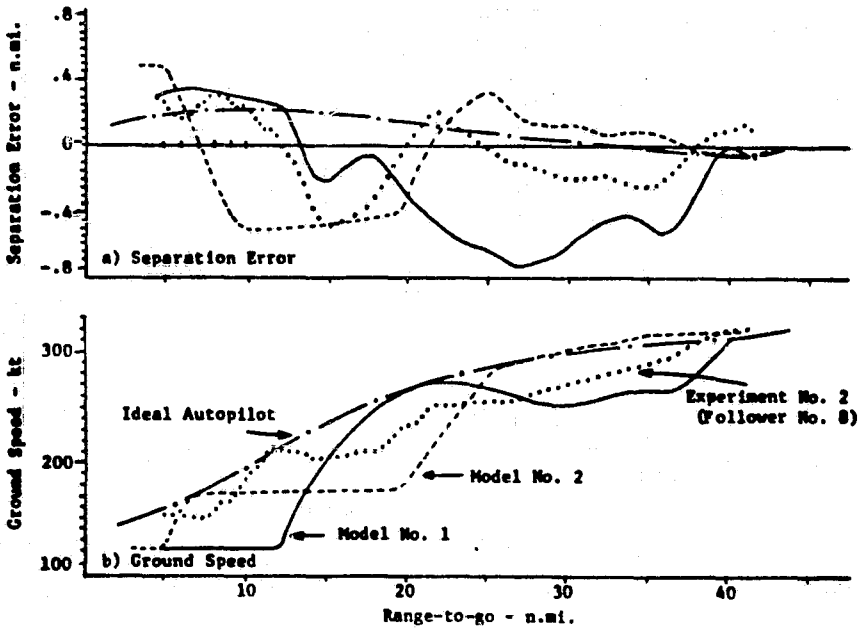


Figure 9. Comparison of Last Aircraft Profiles in Nine-Aircraft String.

(3) Both in separation error and ground speed, the ninth pilots' results from Experiment No. 2 fall between the results predicted by Model Nos. 1 and 2. This is good verification of the models' adequacy in predicting flight system dynamic performance in a string when using the separation criterion expressed by Eq. (3).

(4) Despite the differences seen between the experimental-based models and the autopilot model, the separation errors are acceptable and within 15% of the value specified by Eq. (3). Also, there is no gradual buildup or oscillation of these errors. Thus, we can conclude that the pilot does not induce instability into the string for this aircraft/CDTI configuration.

ADDITIONAL WORK

The above results are preliminary and somewhat ideal. Currently, we are analyzing another set of in-trail following task experimental results based on using the NASA Ames 747 cockpit simulator [15]. This experiment is different from the previously discussed Langley experiments in that (a) the simulated flight begins during the final portion of cruise, (b) the initial separation errors are varied, (c) the vertical control task to follow the desired profile descent is not as automatic as for the TCV simulator, and (d) the lead aircraft have altitude and speed errors in their descent profiles.

There are several more items that should be investigated regarding the in-trail following task. These include the effects of (a) mixed types of aircraft, (b) different separation criteria, (c) winds, (d) some aircraft not being CDTI equipped, and (e) the CDTI sensor and display errors. Beyond this, the stability and dynamic phenomena associated with merging several aircraft into a common string requires a combination of analytical and experimental study. Finally, the dynamic aspects of pilot/air traffic controller interaction for terminal area merging and spacing using CDTI concepts will require investigation.

REFERENCES

1. Anon., "Future Needs and Opportunities in the Air Traffic Control Systems," Report prepared by the Subcommittee on Transportation, Aviation, and Weather, of the Committee on Science and Technology, U.S. House of Representatives, Ninety-fifth Congress, November 1977.

2. Anon., "Joint NASA/FAA Program: Cockpit Display of Traffic Information (CDTI)," Program plan, FAA/NASA, Washington, D.C., February 1980.

3. Sorensen, J.A., and Goka, T., "Analysis of In-Trail Following Dynamics of CDTI-Equipped Aircraft," Report No. 81-9 , Analytical Mechanics Assoc. Inc., Mountain View, CA, June 1981.

4. Kelly, J.R., "Preliminary Evaluation of Time and Distance Spacing Cues Using a Cockpit Displayed Target," NASA TM 81794, August 1980.

5. Palmer, E.A., et al, "Perception of Horizontal Aircraft Separation on a Cockpit Display of Traffic Information," Human Factors, Vol. 22, No. 5., October 1980.

6. Pipes, L.A., "An Operational Analysis of Traffic Dynamics," Jour. of Applied Physics, Vol. 24, No. 3, March 1953.

7. Haight, F.A., Mathematical Theories of Traffic Flow, Academic Press, New York, NY, 1963.

8. Bekey, G.A., "Control Theoretic Models of Human Drivers in Car Following," Human Factors, Vol. 19, No. 4, 1977.

9. Levine, W.S., and Athans, M., "On The Optimum Error Regulation of a String of Moving Vehicles," IEEE Trans. on Automatic Control, Vol. AC-11, No. 3, 1966.

10. Athans, M., and Porter, L.W., "An Approach to Semi-Automated Optimal Scheduling and Handling Strategies for Air Traffic Control," Proceedings of the Joint Automatic Control Conf., St. Louis, 1971.

11. Tobias, L., "Automated Aircraft Scheduling Methods in the Near Terminal Area," J. of Aircraft, Vol. 9. No. 8, August 1972.

11. McRuer, D.T., et al, "A Theory of Human Error," Tech. Report No. 1156-1, Systems Technology, Inc., Hawthorne, CA, May 1980.

12. Baron, S., and Levison, W.H., "The Optimal Control Model - Status and Future Directions, IEEE Decision and Control Conf., 1980.

14. Cavelli, D., "Discrete-Tune Pilot Model," Proc. of the 16th Annual Conf. on Manual Control, Cambridge, MA, May 1980.

15. Chappell, S.L., and Palmer, E.A., "In-Trail Following During Profile Descretes with a Cockpit Display of Traffic Information," Aviation Psychology Conf., Columbus, OH, May 1981.

COMPARISON OF CLOSED LOOP MODEL WITH FLIGHT TEST RESULTS

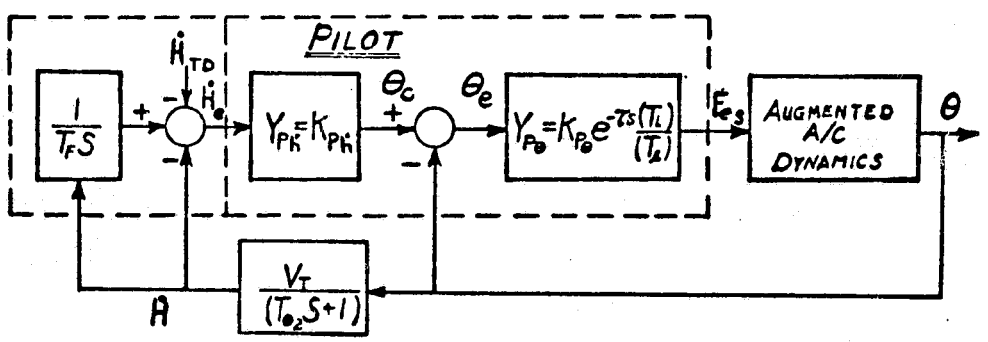
Frank L. George
Air Force Wright Aeronautical Laboratories

Introduction

The analysis presented here is part of a larger effort to develop an analytic technique capable of predicting the landing characteristics of proposed aircraft configurations in the early stages of design. In this first analysis, a linear pilot-aircraft closed loop model is evaluated using experimental data generated with the NT-33 variable stability in-flight simulator. The pilot dynamics are modeled as inner and outer servo loop closures around aircraft pitch attitude, and altitude rate-of-change respectively. The landing flare maneuver is of particular interest as recent experience with military and other highly augmented vehicles has shown this task to be relatively demanding, and potentially a critical design point. A unique feature of the pilot model used here is the incorporation of an internal model of the pilot's desired flight path for the flare maneuver.

Model Development

Data from Ref. 1 suggests the landing flare maneuver can be modeled as a closed-loop tracking task, as pictured in the following sketch.



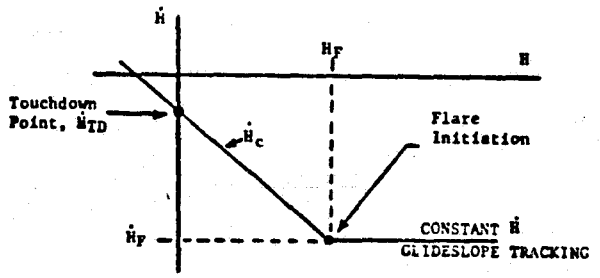
Everything inside the dotted box in this model pertains to the pilot. The pilot model is partitioned into two parts. The left hand part represents the pilot's internal model of the flare maneuver, and provides the basic error signal for tracking. This part of the model will be discussed more in following paragraphs. The right hand part of the model includes all the pilot dynamics which, as will be discussed below, consist of two gains and a time delay, plus possible first order lead or lag compensation.

As can be seen from the sketch, this model only considers pitch axis dynamics and assumes linearized small perturbation equations. The following

variables are pertinent to the discussion.

θ	- rad.	Aircraft Pitch Attitude
$\dot{H}$	- m/sec.	Altitude rate of change, or vertical velocity of aircraft
$\dot{H}_{TD}$	- m/sec.	The pilot's desired vertical velocity at touch-down
$\dot{H}_e$	- m/sec.	Error in desired vertical velocity as perceived by pilot
$1/T_F$	- 1/sec.	Flare model inverse time constant
$K_{p_{\dot{h}}}$	- rad sec/m	Pilot's internal gain to convert $\dot{h}_e$ to an attitude command
θ_c	- rad	Pilot's internal estimate of attitude required to correct $\dot{h}_e$
θ_e	- rad	Error between desired and actual aircraft attitude; as perceived by pilot
K_{p_θ}	- m/rad.	Pilot's internal gain to convert θ_e to an elevator (pitch) control motion
τ	- sec.	Pilot's delay time; approximates internal processing plus neuromuscular delays
T_L	- sec.	Pilot's optional lead compensation time
T_1	- sec.	Pilot's optional lag compensation time
F_{e_s}	- N	Pilot's pitch control force

As suggested in Ref. 1, the pilot's internal flare model, mentioned above, can be represented as a first order exponential. In other words, as a linear relationship between $\dot{H}$ and H, depicted in the following phase plane sketch,



define the pilot's desired sink rate along the flare trajectory as $\dot{H}_c$, where

$$\dot{H}_c = 1/T_F \, H + \dot{H}_{TD}$$

and $1/T_F$, the inverse time constant of the corresponding exponential decay is:

$$1/T_F = \frac{\Delta\dot{H}}{\Delta H} = \frac{\dot{H}_{TD} - \dot{H}_F}{0 - H_F} .$$

Data from Ref. 2 will be used to test the model described above. For the experiment described in Ref. 2, characteristics representative of modern fighter aircraft with augmented dynamics were modeled in the variable stability NT-33. A series of ILS approaches was flown using a nominal glideslope of .044 rad (2.5 degrees); each approach was concluded with a visual flare and landing task. The flare maneuver was initiated at a nominal altitude of 15.24m (50 feet) above ground. A sink rate of .762 m/sec. (2.5 ft/sec.) at touchdown is considered acceptable. The configuration designated as 2-7 in Ref. 2 will be used as a specific example for further discussion here. For the visual flare, it is assumed the pilot has selected a desired touchdown point on the runway so the important speed is equivalent ground speed, which for this example was 61.73 m/sec. (120 knots or 202 ft/sec.) at flare initiation. Using the definitions above, these values give the ideal flare model as,

$$\dot{H} = -.11(H + 6.92) \quad \text{m/sec.}$$

with the exponential solution,

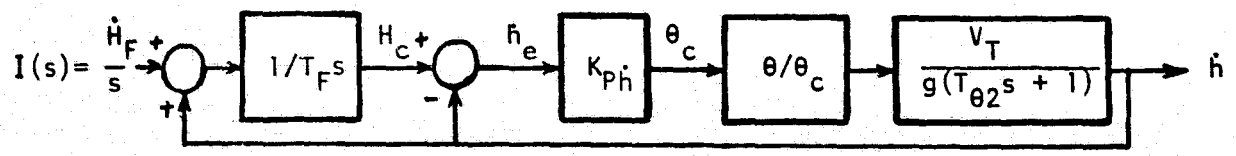
$$H = 22.16(\bar{e}^{.11t} - .31) \quad \text{m}$$

The airplane, actuator, augmentation and feel system dynamics for configuration 2-7 of Ref. 2 were examined and it was determined that a reduced order model using only the short period and command augmentation dynamics plus feel system gain would be adequate in the frequency range of $.1 \leq j\omega \leq 20.$ rad/sec.

The final closed loop model was generated by first closing an inner loop for attitude control, and then an outer loop for altitude rate. Ref. 3 suggests bandwidth and phase angle criteria for attitude control during landing of

$$\omega_B = 1.2 \text{ rad/sec}, \quad \phi_m|_{\omega_B} = \pi/2$$

with a pure time delay of 0.3 sec. For this analysis a pilot time delay of 0.33 sec was used since this value gave a convenient first order Pade' approximate expression for linear frequency domain analysis. The pilot's time delay and any compensation introduced were lumped in the attitude loop closure. Parameters were selected for the inner loop to approach the criteria mentioned above as closely as possible, while maintaining reasonable closed loop damping, for example, $\zeta_{CL} = 0.5$. For the configuration selected, this rationale resulted in a fairly low-gain system, but it was felt this would be consistent with the pilot's desire to not overcontrol so close to the ground. Figure 1 summarizes the characteristics of the attitude closed loop with no pilot compensation considered. Continuing with the outer loop closure, and accounting for initial conditions as suggested in Ref. 1, the final closed loop flare model was,



where the lower case letters are used to indicate perturbation values. The final closed loop transfer function was,

$$\frac{\dot{h}}{\dot{H}_F} = \frac{-38.61\ (S - 6)}{(S + 1.164)(S + 6.005)[S + .771 \pm j\ 1.431][S + 8.356 \pm j\ 8.611]} .$$

The forcing function for this model, using the conditions defined above, was a step scaled to represent the altitude rate at flare initiation.

Model Evaluation

Figure 2 summarizes the response characteristics of the final closed loop system, and Figure 3 compares the landing trajectories of the piloted closed loop model with an ideal exponential curve. The trajectories in Figure 3 representing the piloted system with lead and lag were generated from models developed by the same procedure as described above. The pilot lead and lag were introduced as cascade compensation within the pitch attitude loop. Figure 3 shows the pure gain and lead compensated models both result in early flares and overshoot the ideal exponential touchdown time, leading to long gliding landings which pilots might describe as "floaters". The lag compensated model, on the other hand, generally follows the exponential, but with some low frequency oscillation. An example flare trajectory from the inflight experiment is also shown on Figure 3. As the figure shows, the real flight trajectory does not follow the exponential or the analytic model path very closely except in the last 4 seconds before touchdown, where the flight trajectory is somewhat similar to the exponential. Figure 4 presents a second comparison of the flight path characteristics, using the $\dot{H}$ vs H phase plane. This presentation clearly shows the oscillatory nature of the lag compensated closed loop model. The figure also clearly shows the difference between the actual flight trajectory and the ideal exponential.

Planned Work

Further analysis of the inflight data is planned to investigate differences between the experimental and predicted flare trajectories. It is felt the closed loop model is a valid approach for the landing flare analysis because of pilot comments and the nature of observed pilot control inputs during the flare. Figure 5 shows some examples of the pilot's longitudinal control activity during flare. The experimental data will be examined to determine if significant time-varying or nonlinear characteristics are present which would account for the discrepancies with the linear, constant coefficient model.

Reference 4 investigated STOL aircraft approach and landing using the optimal control pilot model. The flare maneuver was treated as a time opti-

mal problem in that study, resulting in time varying control. An example flare trajectory from the Reference 4 results is plotted in Figure 4. Although the trajectory is rotated because of the steeper approach angle used in the STOL study, the shape generally resembles the NT-33 actual landing trajectory. This observation lends credence to the plan to extend the present investigation to consider time-varying closed loop models.

References

1. Hoh, Roger, Craig, S.J. and Ashkenas, I.L., "Identification of Minimum Acceptable Characteristics for Manual STOL Flight Path Control, Volume III Detailed Analyses and Tested Vehicle Characteristics", US DOT Report No. FAA-RD-75, 123, III, June 1976, DTIC: ADA029250.

2. Smith, Rogers E., "Effects of Control System Dynamics on Fighter Approach and Landing Longitudinal Flying Qualities, Volume I", Air Force Flight Dynamics Lab Report, AFFDL-TR-78-122, March 1978, DTIC: ADA067550.

3. Chalk, C.R., et al, "Revisions to MIL-F-8785B(ASG) Proposed by Cornell Aeronautical Laboratory Under Contract F33615-71-C-1254", Air Force Flight Dynamics Lab Report, AFFDL-TR-72-41, April 1973, DTIC: AD778489.

4. Kleinman, D.L. and Killingsworth, W.R., "A Predictive Pilot Model for STOL Aircraft Landing", NASA Langley Research Center Report, NASA CR-2374, March 1974.

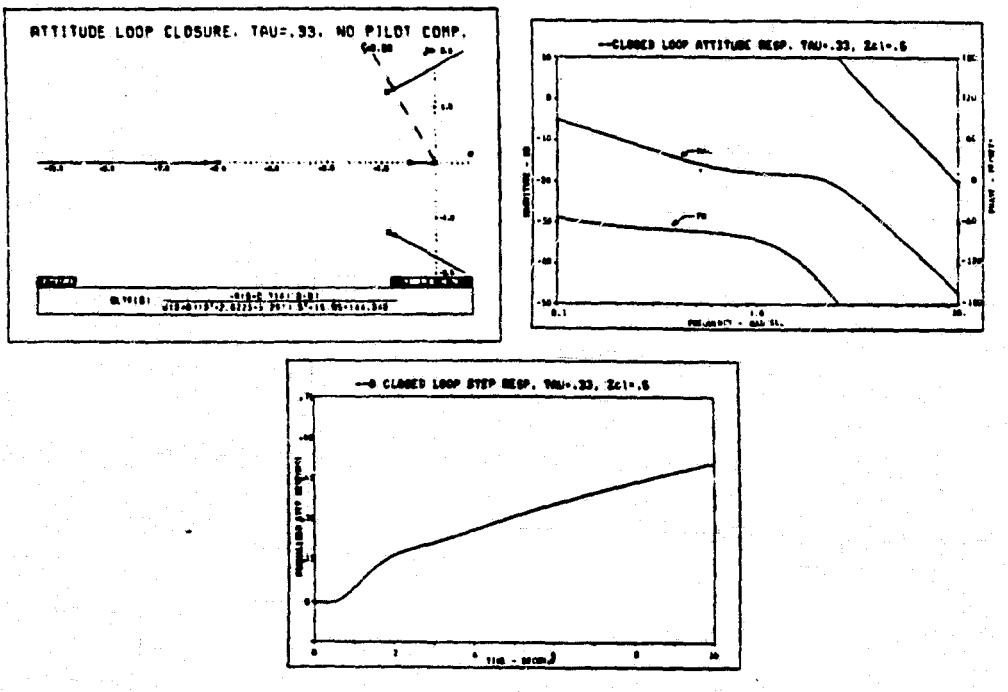


FIGURE 1. Closed Loop Attitude Response with Pure Gain Pilot

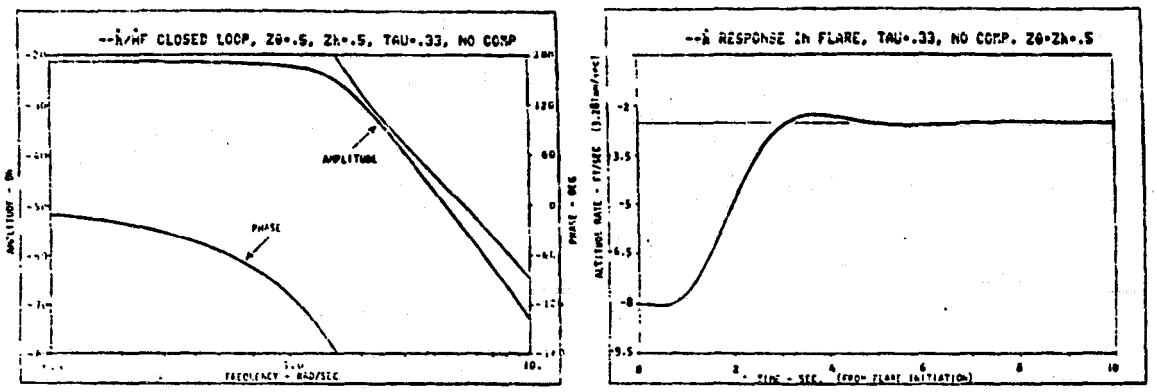


FIGURE 2. Closed Loop Flare Response with Pure Gain Pilot

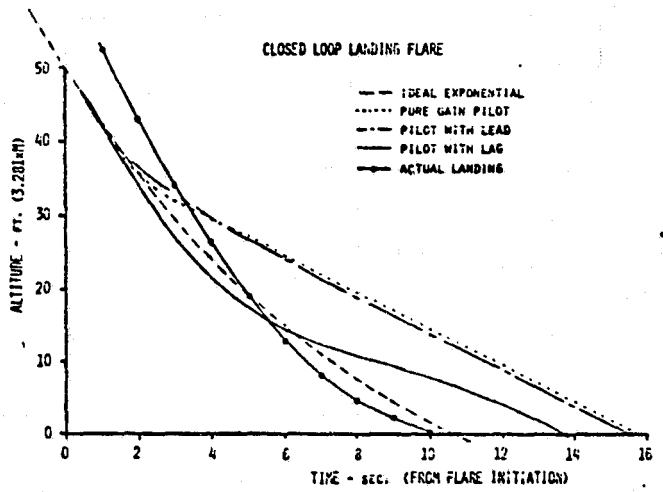


FIGURE 3. Closed Loop Flare Trajectories

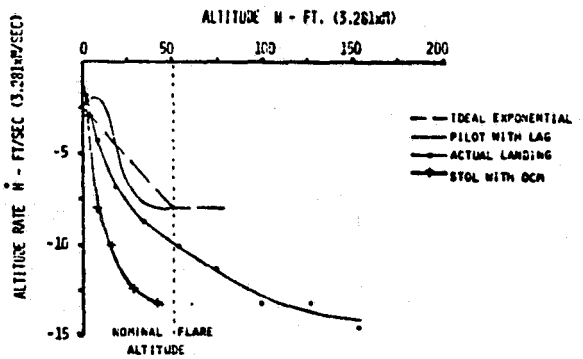


FIGURE 4. Flare Trajectory Phase Plots

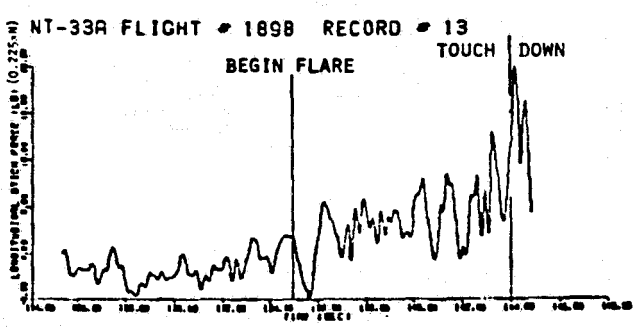


FIGURE 5. Pilot Control Inputs in Flare

FUNCTIONAL STRUCTURE AND DYNAMICS OF THE HUMAN NERVOUS SYSTEM

by: Joseph A Lawrence

This report is an interim report on the status of an effort to define the directions we need to take in extending pilot models. We need these models to perform closed-loop (man-in-the-loop) feedback flight control system designs and to develop cockpit display requirements. The approach we have taken is to first develop a hypothetical working model of the human nervous system by reviewing the current literature in neurology and psychology and second to develop a computer model of this hypothetical working model.

One way of analyzing the nervous system, is to relate the neurons as to their sensory input (Afferent system), processing (Association areas) and output (Efferent system) functions. Konorski, 1965, uses this approach to arrive at some general characteristics of these functional systems.

Konorski organizes the sensory pathways into seven basic functional areas which he calls analyzers: the Visual, Acoustic, Kinesthetic, Somesthetic and Gustatory, Vestibular, Emotional and Olfactory analyzers. These analyzers are so grouped because of similarity in structure and function.

The analyzers have many characteristics in common. Each is built as an n level structure, with each level being an aggregate of neurons performing some signal processing and association tasks. Each level contains cells, "units", that respond maximally to a particular "meaningful" stimulus in the environment. Konorski calls the response a "unitary image response", assuming the nervous system has integrated lower level parts into this meaningful whole. In this manner, the nervous system can operate, for example, from the "i^{th} level on this unitary perception, and/or the unitary perception at this level can act as one piece of a more comprehensive unitary image at the next level. This view is supported by the fact that some outputs from this level proceed to the next higher level while others called "exit" neurons leave the analyzer to subserve lower level reflex loops. The unitary images of higher levels tend to be more complex and refined relative to the level of detail of the image.

Exit neurons proceed into areas we will call association areas. In these areas unitary perceptions between analyzers are associated with one another and with meaningful other perceptions formed over time by logic, thought and concept forming transformations. In these areas, internal models of reality are formed which can be thought of as aggregates of expectation on how present sensory inputs will transform as the result of temperal and spacial movements. Perceptions and models are operated on by logic operation at the higher levels, forming more abstract models, principles and rules etc. Basic drives are associated with remembered past outcomes of patterns of movement to form goals or patterns of planned activity matched to expected outcome. Often repeated patterns of thought or behavior become programmed at many levels as schema and subschema. The models, goals and schema thus become a multi-level hierarchical set. This arrangement is

consistent with philogenetic and ontogenetic theories of nervous system development and function.

The third area by type of signal flow is the efferent systems or output signal generating systems. I will limit my comments to those efferents pertaining to body movement control. Although still somewhat controversial, many believe that output activity is initiated by the formation of an image of expected and desired movement. The image, however, must be accompanied by a desire to move or we experience only the image. The image is the vehicle through which the nervous system is configured for the action. One way for this to occur naturally is for the efferent systems to use the multi-leveled unitary images already created in the Kinesthetic and Somesthetic analyzers. The many exit neurons at each level of the analyzers can subserve this function. What we therefore have as a final output signal is many signals that are somewhat integrated but with the final integration taking place in the spinal column in the areas surrounding the particular motor neurons that drive the (associated) muscles.

There are four additional aspects of the nervous system that are important to a modeling process. One is the term consciousness. Defining conciousness gets us into the whole Mind-Body issue (Spiritual - Physical). I will not try to solve this problem here. But I think we all have the experience of being aware of what we are doing in different ways. Sometimes we can be concentrating on performing a task and are aware of minute details of our goal, our actions and the result of our actions. At other times we are almost observers of our activity especially in well learned tasks. Finally, there are sometimes when we realize that we've driven home almost automatically while our conscious attention was focused on a work or social problem or event. My point is that some tasks are performed with conscious awareness and/or control while others, or parts of some tasks, are performed automatically.

Observing epileptic patients lead Dr. Penfield (1975), a noted neurosurgeon responsible for mapping the functions of a large part of the human cortex, to propose that there are two somewhat independent integrating systems in the human nervous system. One he calls the "Highest Brain Mechanism" (HBM), which, he contends, supports the process of consciousness. The other he calls the "Automatic Sensory-Motor Mechanism", (ASMM), and claims that this mechanism can carry out very complex learned behavior patterns even after the HBM has been immobilized by a petit-mal epileptic seizure.

Whether we accept Dr. Penfield's proposed concept or our own observation of our behavior, we must deal with the job of allocating task assignment to conscious awareness and control,automatic control with conscious awareness or totally automatic control and awareness. But here we must deal with an additional variable.

Consciousness tends to be serial in nature. The automatic system, however, is structured as a hierarchical multi-leveled control system with feedback and feedforward loops and special function generators. The automatic system can and does perform many functions in parallel.

How then can we properly allocate attention to tasks as we study workload? I don't have the answer to this question yet, but it does allow me to appreciate the confusion that presently exists in workload studies. Multi-task modeling efforts will have to deal with finding a workable solution.

The final element is the general state of arousal that the nervous system is in. The nervous system generates a bias signal on the neurons which conditions them for firing. Many factors affect the arousal level. When at low levels, the neurons can't fire with normal signal inputs from processing units. With high levels, the cells fire easily and the nervous system loses its ability to differentiate between signals, thus causing confusion and errors. This factor is important enough to be included in modeling considerations.

In summary then, we can view the nervous system as a multi-level hierarchical automatic control system with parallel processing, feedback and feedforward loops and with the potential of our being consciously aware of its activity and to have conscious control of many functions. Consciousness can be viewed as a master controller acting through the control of the flow of the complex imagery which drives the automatic system.

Reality consists of internal models, remembered transforms of perception, which give a person a sense of understanding himself and his environment, and the ability to establish a set of expectations on what is going to happen next. Indeed, we tend to operate continuously from this set of expectations and are surprised when something different happens. This set of expectations gives us the feedforward capabilities that we use to plan actions to accomplish goals, instantaneous or future. The particular set of expectations available at any moment depends on the motivational state of the nervous system, a very dynamic and difficult variable to consciously control - as any person attempting meditation has already discovered.

A key point to understand is that the set of expectations in the form of images, act to configure the total nervous system as an operating system - both input and output. Because of sensory filtering, our expectations determine to a large degree what we perceive.

Because of the automatic parallel processing, it is possible for several tasks to be performed simultaneously as long as sensory, processing and output elements are not in conflict in performing these functions. In this way, for example, we can maintain our sense of orientation (model) even though we are focused on giving concentrated effort to identifying a target.

With all these variables and this complexity, the multi-task pilot model will have to also be complex and large. Knowing the structure and functions of the nervous system will allow us to make intellegent decisions on how to limit the model, set up experiments and analyze the results.

REFERENCES

Asanuma, H., et al.(1979) Integration in the Nervous System. Tokyo: Igaku-Shoin.

Bandler, R. and Grinder, J. (1975) The Structure of Magic, Vol. 1. California: Science and Behavior Books.

Griffin, D.R. (1976) The Question of Animal Awareness: Evolutionary Continuity of Mental Experience. New York: Rockefeller University Press.

James, W. (1918) The Principles of Psychology. New York: Henry Holt and Company.

Jerison, H.J. (1973) Evolution of the Brain and Intelligence. New York: Academic Press.

Konorski, J. (1970) Integrative Activity of the Brain: An Interdisciplinary Approach. Chicago and London: The University of Chicago Press.

Oakley, D.A. (1979) Cerebral cortex and adaptive behavior. IN Oakley, D.A. and Plotkin, H.C. (1979) Brain, Behaviour and Evolution. New York: Methuen.

Penfield, W. (1975) The Mystery of the Mind: A Critical Study of Consciousness and the Human Brain. New Jersey: Princeton University Press.

Sommerhoff, G. (1974) Logic of the Living Brain. New York: John Wiley and Sons.

Sperry, R.W. (1979) Consciousness, freewill and personal identity. IN Oakley, D.A. and Plotkin, H.C. (1979) Brain, Behaviour and Evolution. New York: Methuen.

Yho, C.H. (1979) The anatomy of the vertebrate nervous system: an evolutionary and developmental perspective. IN Oakley, D.A. and Plotkin, H.C. (1979) Brain, Behaviour and Evolution. New York: Methuen.

28

ON THE USE OF THE OCM'S QUADRATIC OBJECTIVE FUNCTION AS A PILOT RATING METRIC

David K. Schmidt

School of Aeronautics and Astronautics
Purdue University
West Lafayette, IN 47907

In several previous works, (Refs. 1-3) a correlation between the magnitude of the quadratic objective function from an optimal control pilot model and the subjective rating of the vehicle and task has been discussed. Since such a correlation would provide a valuable tool for handling qualities research and flight control synthesis, as used in Reference 4 and 5 for example, validating it over a wide range of tasks and plant dynamics is appropriate.

To this end, an anlysis of Arnold's (Ref. 6) simulation results for fourteen aircraft configurations flight tested earlier by Neal and Smith (Ref. 7) has been completed. A fixed set of pilot model parameters, given in Table 1, were found for all cases in modeling the simulated regulation task. The agreement obtained between performance statistics is shown in Figure 1, and a strong correlation, shown in Figure 2, was obtained between the cost function and rating. Furthermore, modeling the same fourteen configurations in the tracking task used by Neal and Smith indicated reasonable correlation as well, considering no experimental data is available from the Neal and Smith tests to check the pilot model parameters in this case.

However, when evaluating other configurations tested by Neal and Smith that included higher-order control system dynamics, the pilot rating/cost magnitude sensitivity, or the slope of the regression, appeared to be greater. This is indicated in Figure 3.

All these configurations have identical short period eigenvalues (which would yield Level 1 ratings according the the Mil Spec. 8585B), and yet ratings as high as eight were obtained in the flight tests due to the other dynamic modal characteristics.

The significant factors are that correlation between pilot rating and cost function is evident for these cases, but the sensitivity (slope) of the rating would seem to be much greater than that exhibited in the previous figure . The reasons for this apparent difference in sensitivity are

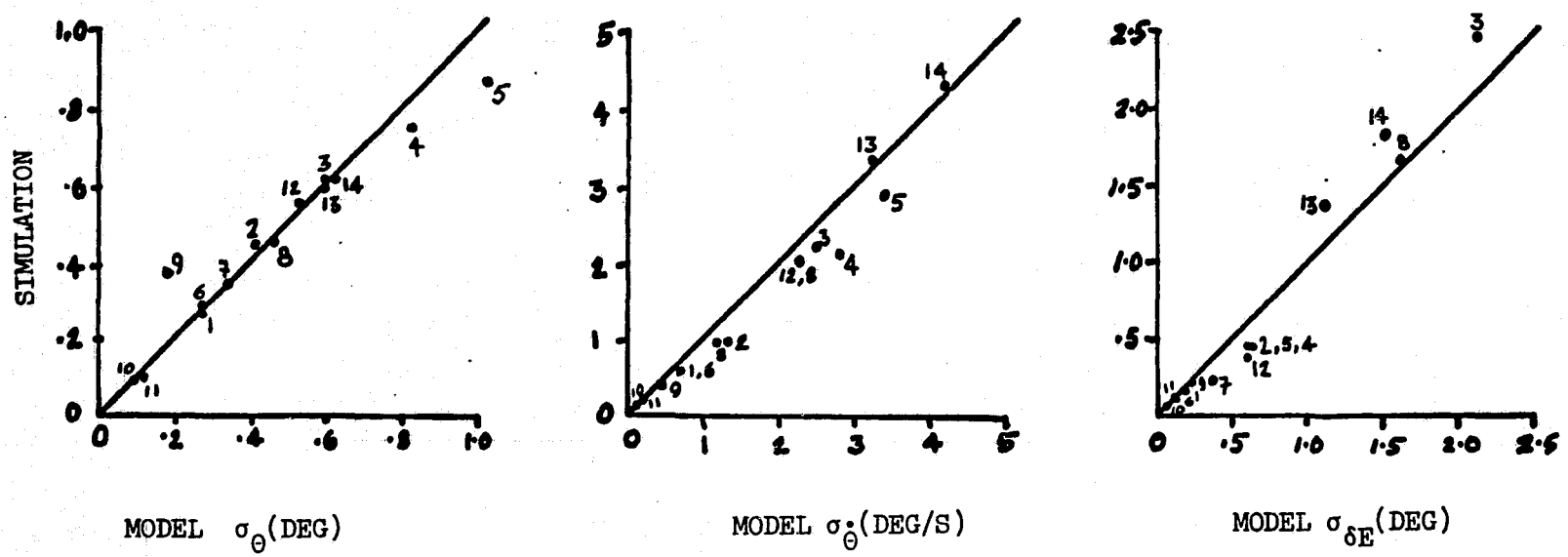


FIG. 1, CORRELATION BETWEEN ARNOLD'S SIMULATION AND MODEL RMS PERFORMANCE

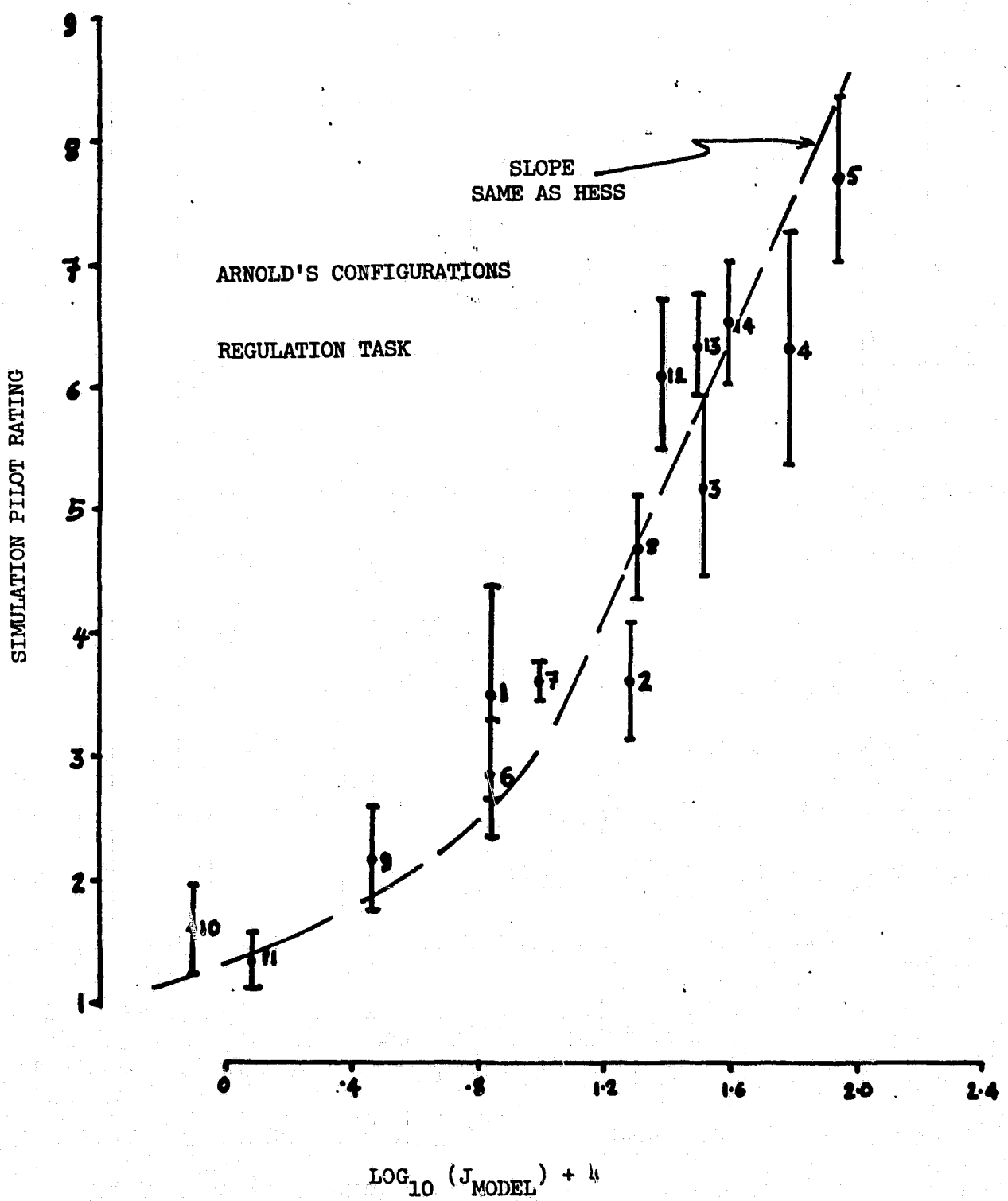


FIG. 2, OBJECTIVE FUNCTION/PILOT RATING CORRELATION

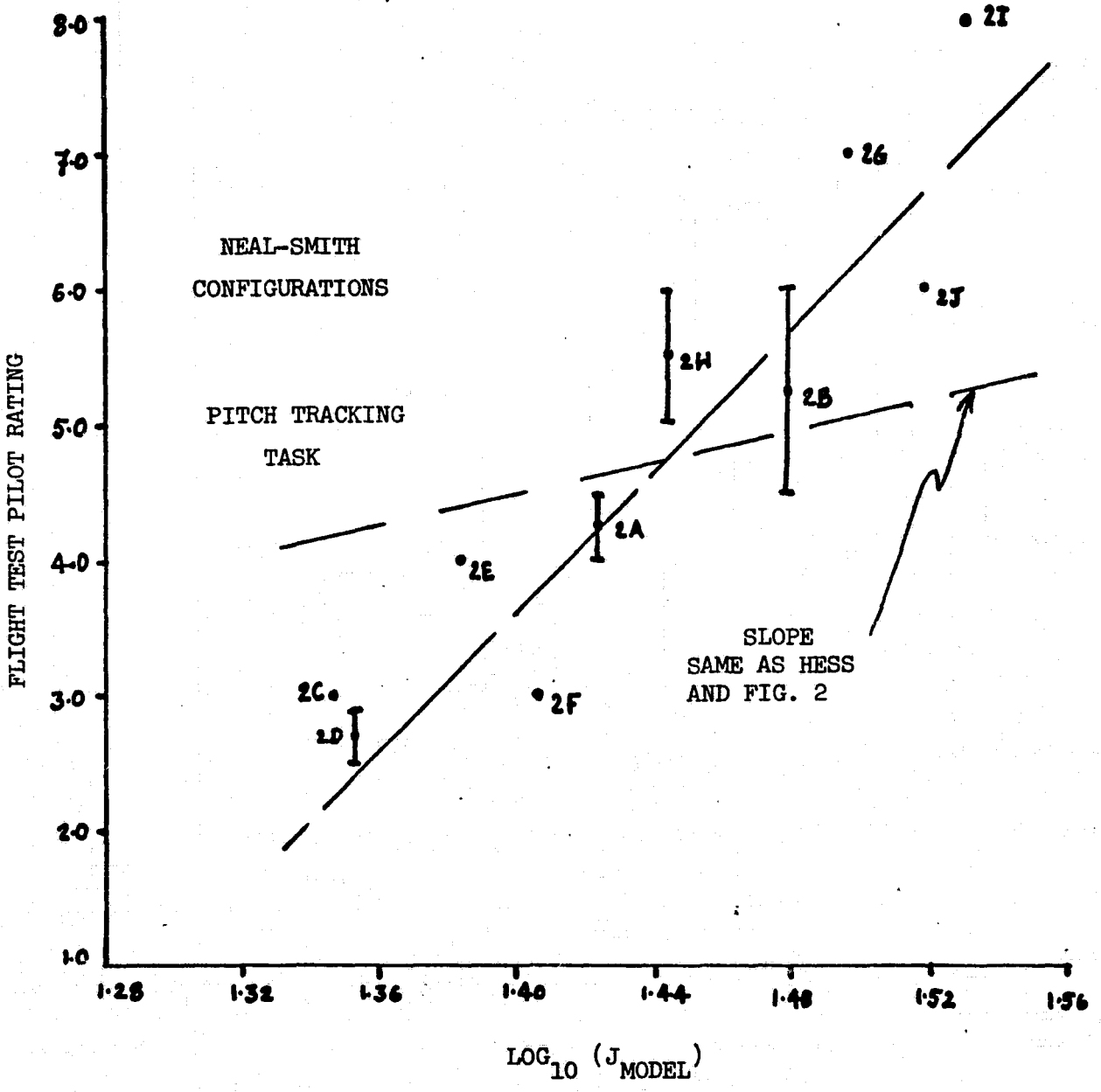


FIG. 3, CORRELATION FOR HIGH-ORDER DYNAMICS CONFIGURATIONS

Table 1

Pilot Model Parameters For Arnold

Simulation Cases

Observation vector,	$\bar{y}_p^T = [\theta_{error}, \dot{\theta}_{error}]$
Objective function weights,	$Q_\theta = 25.0$ $Q_{\dot{\theta}} = 0.1$
Fractional attention allocation,	0.5 on θ and $\dot{\theta}$
Full attention observation noise ratio,	-20dB
Observation thresholds,	$T_\theta = .002°$ $T_{\dot{\theta}} = .004°/sec$
Observation delay,	$\tau = 0.2$ sec
Neuromuscular lag,	$\tau_N = 0.1$ sec
Neuromotor noise ratio,	-20dB
Control input δ_{stick} to minimize	$J_p = E\left\{\lim_{T\to\infty} \frac{1}{T}\int_o^T (Q_\theta\theta^2+Q_{\dot{\theta}}\dot{\theta}^2+g\delta_{st}^2)dt\right\}$

open to conjecture at this point, but the following are put forth as possibilities.

1) The sensitivity is greater for dynamics significantly different from "rigid-body-only" dynamics, with which the pilot has more familiarity. Just as rating sensitivity is often higher for greater task difficulty, the significantly different dynamic characteristics may lead to more sensitive ratings from the pilots.

2) Or, the sensitivity is not really different from that shown previously, but the pilot model is incorrect for these aircraft and hence the cost function is not correct. Note that in the absence of rms statistics from the experimental work, we are not really confident that the pilot model, calibrated from simulation results on low-order dynamics, is correct. Even possible too is that the OCM of the pilot may need modification when investigating higher order dynamic systems. At any rate, the effect of pilot model inaccuracies may be the prediction of an incorrect trend or sensitivity here.

It appears that when significant higher-order dynamics are due to aeroelastic (or other low damped) modes, the latter hypothesis can be supported.

To be considered are the fixed-base simulation results of Yen (Ref. 8), in which a B-1-type vehicle was evaluated in an attitude tracking task with a very low-frequency discrete command signal. Three cases (given below) of vehicle dynamics are considered, each including short-period, phugoid, and two aeroelastic modes.

Table 2

Three Cases of Vehicle Dynamics

	ζ_{sp}	ω_{sp} rad/sec	ζ_{ph}	ω_{ph} rad/sec	ζ_{1e}	ω_{1e} rad/sec	ζ_{2e}	ω_{2e} rad/sec
1.	0.5339	2.806	0.0197	0.0708	0.0494	13.312	0.0215	21.354
2.	0.5235	2.572	-0.0006	0.0573	0.0877	8.789	0.0213	21.356
			Real Roots					
3.	0.5217	1.769	+0.0910		0.1999	5.866	0.0213	21.357
			-0.0767					

Now the simulated and modeled (via OCM) tracking error and pilot rating are given in Figures 4 and 5. Specifically note the results for case 3, in which the phugoid is unstable and the first elastic mode frequency is very low.

Now the total attitude angle observed by the pilot is the sum of the rigid, or mean axis attitude plus the change in local attitude due to structural flexure, or

$$\Theta_{Total} = \Theta_{Rigid} + \Theta_{Elastic}$$

If it is assumed that the pilot is attempting to regulate total attitude error given by

$$\varepsilon_T = \Theta_{Command} - \Theta_{Total}$$

the tracking error obtained from the model is significantly less (1.9°) than obtained in simulation.

However, if the rigid, or mean-axis error is assumed to be regulated, rather than total error, the results agree extremely well! That is, this mean-axis error, given by

$$\varepsilon_{Mean} = \Theta_{Command} - \Theta_{Rigid}$$

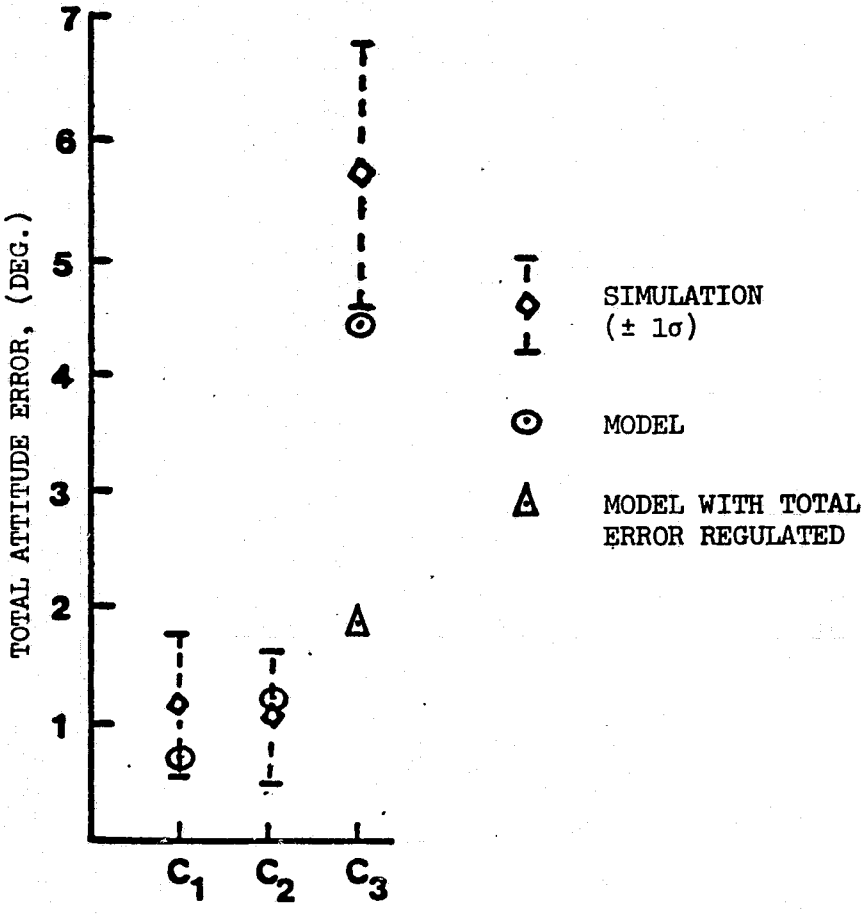


FIG. 4 ATTITUDE ERROR COMPARISON

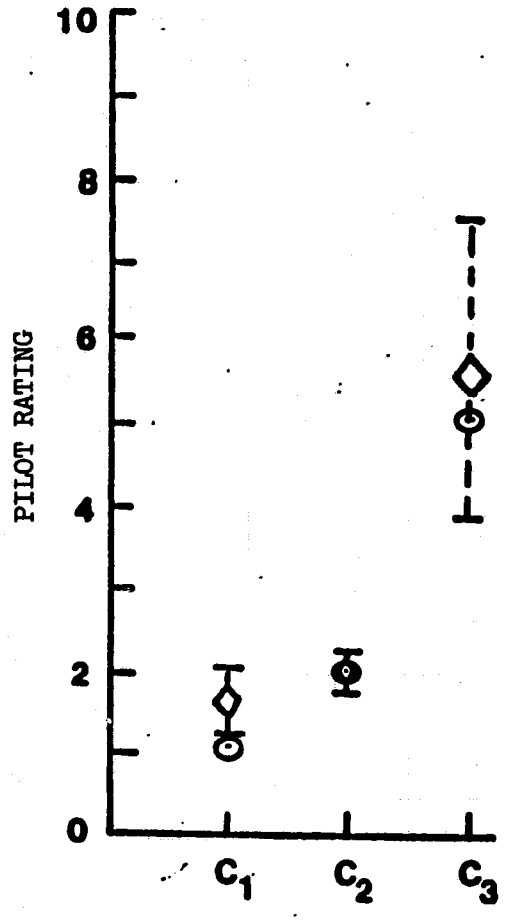


FIG. 5 PILOT RATING COMPARISON

is weighted in the pilot's objective function rather than total error, and since total attitude and total error is what is being observed by the pilot, he must estimate ϵ_{Mean} and then attempt to minimize the estimate. Finally, the pilot rating results are obtained by using this different objective function with the original sensitivity (slope) from Figure 2!

Based on this approach, it appears that rating sensitivity is constant, and that the degredation in rating and performance for case 3 may be primarily due to the difficulty in estimating the mean-axis error when the aeroelastic mode frequency approaches that of the "rigid-body" or mean-axis short period frequency. Furthermore, this approach is in contrast to that of Swaim, (Ref. 9) where the assumption of a reduced-order pilot model is made.

ACKNOWLEDGEMENTS

Discussions with Dr. Swaim have been most helpful in developing the ideas presented herein.

Appreciation for the numerical results presented is expressed to Mr. S. N. Prasad, Ph.D. candidate, School of Aeronautics and Astronautics, Purdue University.

Portions of this work were supported by the NASA Dryden Research Center under grant number NAG4-1.

REFERENCES

1) Hess, R.A., "A Method for Generating Numerical Pilot Opinion Ratings Using the Optimal Pilot Model", NASA TMX-73, 101, 1976.

2) Hess, R.A., "Prediction of Pilot Opinion Ratings Using an Optimal Control Pilot Model", Human Factors, pp. 459-475, 1977.

3) Hess, R.A., "A Pilot Modeling Technique for Handling Qualities Research", AIAA Paper 80-1624, Atmospheric Flt. Mech. Conf., August 1981.

4) Schmidt, D.K., "Optimal Flight Control Synthesis Via Pilot Modeling", AIAA J. of Guid. and Cont., Vol. 2, No. 4, 1979.

5) Schmidt, D.K. "Pilot-Optimal Augmentation of the Air-to-Air Tracking Task", AIAA J. of Guid. and Cont., Vol. 3, No. 5, 1980.

6) Arnold, J.D., "An Improved Method of Predicting Aircraft Longitudinal Handling Qualities Based on the Minimum Pilot Rating Concept", AFIT Thesis, GGC/MA/73-1, 1973.

7) Neal, T.P. and Smith, R.E., "An In-Flight Investigation to Develop Control System Design Criteria for Fighter Airplanes", Vols. I and II, AFFDL-TR-70-74, 1970.

8) Yen, W.Y., "Effects of Dynamic Aeroelasticity on Handling Qualities and Pilot Rating", Ph.D. Dissertation, Purdue University, West Lafayette, IN, Dec., 1977.

9) Swaim, R.L. and Poopaka, S., "An Analytical Pilot Rating Method For Highly Elastic Aircraft", AIAA Paper 81-1772, Presented at 1981 Guid. and Cont. Conf., Albuquerque, NM, Aug., 1981.

A Quasi-Newton Procedure for Identifying Pilot-Related Parameters of the Optimal Control Model

by

William H. Levison
Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, MA 02238

ABSTRACT

Progress is reported in the development and application of a quasi-Newton gradient search procedure for identifying independent pilot-related parameters of the optimal control model for pilot/vehicle systems. The computational efficiency of the scheme originally implemented by Lancraft and Kleinman has been improved. A sensitivity analysis procedure is described that allows one to determine (1) whether or not a given model parameter is required to match a specific experimental result, and (2) which experimentally-induced parameter changes are "significant"; i.e., required to account for behavioral and performance differences. Application of the identification scheme to training effects in a manual control task is described.

INTRODUCTION

Considerable effort has been expended over nearly four decades to develop mathematical models for predicting and diagnosing human operator response behavior in closed-loop control tasks. Predictive models are desired for the general purpose of extrapolating knowledge, gained from man-in-the-loop studies, to tasks in which experimental data have not yet been obtained. Diagnostic models, on the other hand, are intended to help quantify and interpret the effects of stress, and other aspects of the task environment, on operator response capabilities.

Although the same model form may be used for both prediction and diagnosis, the treatment of independent model parameters is different. For predictive applications, one desires a set of independent model parameters that are either constant or selectable on the basis of well-defined adjustment rules. For diagnostic applications -- particularly when the influence of environmental or task parameters is not well understood -- independent model parameters will often be adjusted to provide the best match to the experimental data, with little constraint on parameter values. For this type of model usage, a well-defined procedure is required to uniquely identify (i.e., quantify) independent model parameters from the experimental data, and to indicate the reliability of the identified parameter values.

Particular impetus for developing reliability metrics for identified parameters has been provided by a study of the effects of learning on the human controller's response strategy [1]. As shown later in this paper, parameter values undergo rather large changes during training, and a determination as to which of these changes are statistically meaningful can provide guidance for the development of models for learning behavior.

This paper describes a procedure for identifying and testing parameters of the "optimal control model" (OCM) for the human operator in steady-state control tasks. Typical independent -- or "pilot-related" -- parameters to be identified from laboratory tracking data are time delay, motor time constant, (equivalently, a "cost" penalty on rate-of-change of control), motor noise covariance, and an observation noise covariance for each perceptual input variable used by the operator. Readers unfamiliar with this model are directed to the recent review article by Baron and Levison [2] and to the references cited therein.

REVIEW OF THE QUASI-NEWTON IDENTIFICATION PROCEDURE

In this section we first review the process of adjusting independent model parameters to provide a best match to the experimental data. We then offer a technique for deriving an analytic approximation to the sensitivity of the matching error to perturbations in these parameters. Finally, certain implementation details are discussed. In the interest of conciseness, only major results are presented here. Derivations have been reported in more detail in [1].

The Basic Minimization Procedure

Consider the task of adjusting model parameters to minimize a scalar matching error $J=\underline{e}'\underline{W}\underline{e}$, where each element e_i of the column vector $\underline{e}$ is the difference between the ith measured data point and the corresponding model prediction, and each element w_i of the diagonal matrix $\underline{W}$ is a weighting coefficient. In a particular application, the matching error J will correspond to a particular choice of parameter values $\underline{p}$. The objective of the search procedure is to find a new parameter set $\underline{p}+\Delta\underline{p}$ such that J is minimized.

To implement the search scheme, we initially assume that model predictions (and, therefore, prediction errors) vary linearly with model parameters. Thus, $\Delta\underline{e} = \underline{Q}'\Delta\underline{p}$, where

$$q(i,j) = \partial e_j/\partial p_i \qquad (1)$$

Solving for minimum J as a function of $\underline{p}$, we obtain

$$\Delta\underline{p} = -[\underline{QWQ}']^{-1}\underline{QWe} \qquad (2)$$

Now, since model input/output relationships are seldom totally linear, two or more iterations of the procedure are required until some convergence criteria are satisfied. In some cases, the parameter change computed as shown in Eq(2) will yield a scalar matching error greater than the starting value. Therefore, it is often useful to augment the minimization procedure described above with a line-search scheme to optimize the magnitude of $\underline{p}$.

Sensitivity Analysis

An indication of parameter estimation reliability can often be obtained through sensitivity analysis relating changes in the scalar matching error to perturbations in model parameters. In general, estimates of parameters that have a high impact on matching error can be considered more reliable than estimates of parameters having a smaller impact.

If model predictions are linear in the parameters, we may analytically derive the sensitivity of the scalar modeling error to perturbations in model parameters about the optimal (best-matching) set. One may compute the sensitivity to a given parameter with the remaining model parameters held fixed, or with remaining parameters reoptimized. We shall compute sensitivity according to the latter definition because, by allowing tradeoffs among parameters in terms of minimizing matching error, it provides a more stringent reliability measure.

The sensitivity of the matching error J to a change in a single parameter p_i is

$$J = \underline{v}'\underline{QWQ}'\underline{v}(\Delta p_i)^2 \qquad (3)$$

where J is the increment in J about its minimum value, p_i the change in parameter p_i about its optimum, and $\underline{v}$ is a column vector that has unity value for the ith element and values for remaining elements given by the following expression:

$$\underline{v}_r = -[\underline{Q}_r\underline{WQ}'_r]^{-1}\underline{Q}_r\underline{W}\underline{q}_i \qquad (4)$$

where $\underline{q}_i = \mathrm{col}(q_{i,1}, q_{i,2}, \ldots)$, and the subscript "r" indicates vectors and matrices with omission of rows and columns corresponding to the ith model parameter. (See [1] for a derivation of this result.) The change in matching error, therefore, varies as the square of the change in parameter value, given the underlying assumption of linearity between model parameters and model predictions.

Implementation of Manual Control Studies

Application of the QN method for analysis of human operator performance in continuous control tasks has been reported by Lancraft and Kleinman [3]. Described below is a revised implementation that was used to perform the model analysis described later in this paper.

Two criteria must be defined in order to apply the identification procedure: (1) a definition of a scalar matching error to be minimized by the QN scheme, and (2) convergence criteria to determine when the minimum modeling error has been approached sufficiently closely to justify termination of the minimization procedure.

Matching error is similar to that used by Lancraft and Kleinman:

$$J = \frac{1}{N_1}\sum_{i=1}^{N_1}\left(\frac{G_i-\hat{G}_i}{\sigma_{G_i}}\right)^2 + \frac{1}{N_2}\sum_{i=1}^{N_2}\left(\frac{P_i-\hat{P}_i}{\sigma_{P_i}}\right)^2 + \frac{1}{N_3}\sum_{i=1}^{N_3}\left(\frac{R_i-\hat{R}_i}{\sigma_{R_i}}\right)^2 + \frac{1}{N_4}\sum_{i=1}^{N_4}\left(\frac{S_i-\hat{S}_i}{\sigma_{S_i}}\right)^2 \quad (5)$$

where N is the number of valid measurements in the jth measurement group; G,P, are the gain (dB) and phase shift (degrees) of the ith describing function point to be matched; R is the corresponding control-stick "remnant" measurement (dB); and S is the ith variance score to be matched (units different for different tracking variables). indicates standard deviation of an experimental data point, and the symbol "^" ("hat") indicates a model prediction.

Inclusion of the experimental deviations in the scalar modeling error allows each error component to be weighted inversely by the reliability of the data. To prevent the matching criterion from giving excessive weights to variables that have very low experimental variability, the following minimum standard deviations are imposed: 0.5 dB for magnitude and remnant, 3 degrees for phase, and 5% for the ensemble mean for variance scores.

Weighting inversely by standard deviation also converts each error term into a dimensionless number, thereby allowing accumulation of matching errors into a single metric. Thus, the matching error defined in Eq(5) approximates the average number of standard deviations of mismatch. A numerical score of J=4 reflects an average modeling error of 1 standard deviation (i.e., an average error of unity per measurement group).

The minimization procedure is terminated when the following conditions jointly obtain for two successive iterations: (1) reduction of the matching error by less than 0.5%, and (2) changes in all identified parameters by less than 2%. The first criterion is based on the fact that the sensitivity of matching error to small perturbations of model

parameters is relatively low in the vicinity of the minimum (a consequence of the quadratic matching error). The second criterion prevents termination resulting from a compensating "overshoot"; i.e., a situation in which successive estimates of one or more parameters bound the optimal values in such a way as to yield essentially the same modeling error.

A number of modifications have been made to the original implementation in order to improve computational efficiency. First, the search is performed on the logarithms of the parameters. This transformation modestly increases the degree of linearity between model parameters and model outputs, and it prevents the assignment of out-of-bounds (i.e., negative) values to parameters during the course of the search. Second, in order to minimize numerical difficulties with inversion of the expression $\underline{QWQ}'$, we omit from the search procedure (i.e., keep fixed), at a given iteration, any parameter having a negligible influence on the matching error. In addition, to reduce the chance of convergence to a local minimum appreciably removed from the global minimum, an individual parameter is allowed to undergo no more than a ten-fold increase or decrease from one iteration to the next.

Finally, a binary section scheme is employed to prevent divergence of the QN scheme due to nonlinear relationships between model inputs and outputs. If necessary, binary section is repeated until (1) matching error is reduced from one iteration to the next, or (2) until four attempts fail to reduce matching error, at which point the minimization scheme is terminated. Further details regarding implementation are documented by Levison [1].

As is true with any numerical search procedure, the probability of convergence to a global minimum is enhanced by the selection of an initial set of model parameters that are close to the optimal set. The following rules for initializing model parameters appears to yield good results with the QN procedure: (1) cost of control rate such that motor time constant = 0.1 seconds; (2) time delay = 0.2 seconds; (3) observation noise covariance to achieve a noise/signal ratio of -20 dB for each perceptual variable assumed to be utilized by the operator; and (4) motor noise covariance to achieve a noise signal ratio of -50 dB, normalized with respect to control-rate variance.

SIGNIFICANCE TESTING

In the following discussion we assume that the data base being subjected to model analysis reflects a significant difference in human operator response behavior, as determined by some standard quantitative test for significance. We then wish to test the hypothesis that the various data sets can be modeled by the same set of model parameters. Failure to support this null hypothesis indicates that parameter differences are also significant.

A cross-comparison scheme was developed and tested against data obtained in manual control studies. In general, this method may be employed to provide a qualitative significance test on parameter differences obtained from modeling the results of two experimental conditions. This method employs an empirical sensitivity test as described below.

Assume that model parameters have been identified from two data sets corresponding to, say, the "baseline" and "test" experimental conditions; our task is now to test the null hypothesis that a single set of model parameters provides a near-optimal match to the baseline and test data. To perform this test, we first identify the following three sets of pilot parameters: (1) the set that best matches the baseline data, (2) the set that best matches the test data, and (3) the set that provides the best joint match to the baseline and test data. For convenience, we shall refer to the parameters identified in step 3 as the "average parameter set".

We next compute the following four matching errors:

J(B,B) = matching error obtained from baseline data, using parameters identified from baseline data (i.e., best match to baseline data).
J(B,A) = matching error obtained from baseline data, using average parameter set.
J(T,T) = best match to test data.
J(T,A) = matching error obtained from test data, using average parameter set.

Finally, we compute the following "matching error ratios": MER(B) = J(B,A)/J(B,B), MER(T) = J(T,A)/J(T,T), and, if we wish to reduce the results to a single number, the average of these two error ratios. In a qualitative sense, the greater the matching error ratios, the more significant are the differences between the parameters identified for the baseline and test conditions.

As shown by Levison [1], a good approximation to the joint match to multiple data sets can be obtained by simply matching the average data. Thus, to obtain the "average parameter set", one would first obtain a point-by-point ensemble average of the (reduced) baseline and test data, and then identify parameters to match the average data set. This procedure is valid if the same task description applies to the two experimental conditions; i.e., if both tasks can be modeled identically except for quantitative differences in pilot-related parameters. Experiments designed to explore training effects, environmental stress, or interference from other concurrent tasks often meet this restriction.

In addition to providing a collective test of the entire parameter set, this scheme may also be used to test a single parameter or a subset of parameters. Suppose, for example, one wishes to test apparent differences

in the time delay parameter. The matching errors J(B,B) and J(T,T) would be computed as described above. The errors J(B,A) and J(T,A), however, would be computed with only the time delay parameter fixed at its "average" value; remaining parameters would be re-optimized.

APPLICATION TO STUDIES OF HUMAN OPERATOR PERFORMANCE

Two applications of the cross-comparison scheme for significance testing are illustrated below. First, data from rate- and acceleration-control systems are analyzed to determine the degree of parameterization required in each case. Second, we analyze the effects of training on operator response behavior.

Parameterization Requirements

The QN analysis methodology described above was applied to data obtained from two manual control studies: one utilizing a rate-control system [4], and one using approximate acceleration-control dynamics [5]. In both studies, a pseudo-random forcing function was applied in parallel with the operator's control input, and subjects were trained to near asymptotic levels of performance. The data bases subjected to model analysis were obtained by averaging performance measures from three subjects for the first study, and from eight subjects for the second.

The following five independent model parameters were identified in each case: (1) observation noise on error, (2) observation noise on error rate, (3) pseudo motor noise, (4) time delay, and (5) relative cost of control rate (equivalently, motor time constant). Identification was repeated for each data base with time delay, pseudo motor noise, and rate observation noise omitted individually from the analysis.* When any one parameter was omitted, remaining parameters were re-optimized to yield minimum modeling error.

Matching error ratios were computed by normalizing the scalar modeling error obtained with a parameter omitted, to the modeling error obtained with all five parameters identified. The matching error ratios presented in Table 1a indicate that all three parameters tested were required to parameterize the data obtained from the rate-control system. That is, with any single parameter omitted, the matching error increased by a factor of three or more. Time delay and rate observation noise were also required to match the acceleration-control data, but pseudo motor noise proved to be an extraneous parameter (matching error ratio of 1.02) for this data set.

---

* The mathematical structure of the model requires finite, non-zero values for cost of control rate and (for these tasks) for observation noise on error. Therefore, model analysis was not performed with these parameters omitted.

Table 1. Model Parameterization Requirements

Model Parameter	Rate Control	Acceleration Control
a) Effect of Omitting Parameter on Matching Error Ratio		
Rate Observation Noise	8.6	8.3
Motor Noise	3.6	1.0
Time Delay	65.3	10.7
b) Inverse Sensitivity		
Displacement Obs. Noise	1.8	33.3
Rate Observation Noise	2.1	5.9
Motor Noise	4.7	62.2
Time Delay	0.4	2.3
Cost of Control Rate	2.8	5.3

Table 1b contains the analytic inverse sensitivity computations for each of the five parameters identified in the initial analysis for each data base. These measures, which indicate the decilog change required to increase matching error by 4 units, were computed analytically during the QN search as part of the parameter reduction procedure described earlier.

The analytic sensitivity predictions correlate well with the empirical matching error ratios shown in Table 1a. For a given control system, matching error ratio (a direct measure of sensitivity) varies inversely with predicted inverse sensitivity. In particular, especially large inverse sensitivity is shown for the one parameter (motor noise, acceleration control task) that is considered extraneous. Therefore, the analytic sensitivity prediction provides guidance to required model parameterization.

Three conclusions can be drawn from this illustration. First, the gradient search technique in general, and the QN identification scheme in particular, has the intrinsic capability to identify time delay, motor noise, and rate observation noise -- a capability that has not been demonstrated by maximum likelihood schemes [6]. Second, analytic sensitivity computations performed as part of the QN search procedure provide an indication of the required parameterization. Finally, the ability to identify a particular model parameter will, in general, depend on the specifics of the experimental data base.

A Study of Training-Related Performance Differences

The model analysis scheme described in this paper was used in a recent study to quantify and interpret the effects of training on human operator performance [1]. The experimental data base was obtained from an earlier study which explored the effects of delayed motion cuing on roll-axis tracking performance. Of interest here are pre-transition performance measures obtained from subjects initially trained fixed-base.

Subjects were required to maintain simulated wings-level attitude in a single-axis laboratory tracking task. Vehicle dynamics were representative of a high-performance fighter aircraft in the roll axis, and a zero-mean gust environment was simulated. Except for a brief familiarization period, all training and data trials were conducted with the external forcing function and were digitally recorded for subsequent analysis and modeling. Details of the experiment have been reported by Levison, Lancraft, and Junker, [7].

Frequency response measures are shown in Figure 1 for a single subject very early in training ("Early Training") and for the final pre-transition training session ("Late Training"). This training interval represented about 70 experimental trials. Training induced the following changes in response behavior: an increase in amplitude ratio ("gain") at all frequencies, (2) a decrease in high-frequency phase lag, and (3) a reshaping of the control-stick remnant spectrum to yield decreased remnant power at low frequencies and increased remnant at high frequencies. RMS tracking error (not shown) decreased by almost a factor of two over the course of this training interval.

These gain and phase-shift changes are consistent with improved tracking efficiency. While not obvious, training-related changes in remnant are also indicative of improved tracking efficiency and are consistent with the hypothesis (borne out by model analysis) that training leads to decreased response variability and increased man/machine response bandwidth.

Table 2 shows pilot-related model parameters for two test subjects. Parameters are shown for an average of 2-4 trials very early in training and for the average of the final four trials. (The smooth curves shown in

Table 2. Effects of Training on Pilot-Related Model Parameters

State of Training	Subject	Pilot Parameter				
		P_{ye}	$P_{y\dot{e}}$	P_u	τ	T_n
Early	CP	-5.3	-18.6	-28.2	.230	.343
Late		-21.6	-16.4	-29.3	.162	.169
Early	TB	-11.0	-15.9	-70.1	.198	.162
Late		-21.2	-17.4	-56.9	.219	.121

P_{ye} = error observation noise/signal ratio, dB

$P_{y\dot{e}}$ = error rate observation noise/signal ratio, dB

P_u = motor noise/signal ratio, dB

T_n = motor time constant, seconds

τ = time delay, seconds

Figure 1 are model predictions obtained with the parameter sets shown for Subject CP.) To be consistent with previous publications, the relative weighting coefficient for control rate is shown as an equivalent motor time constant [1], and noise covariances are presented as noise/signal ratios.

The following effects of training are noted: (a) a substantial reduction in the observation noise associated with perception of tracking error, (b) a substantial reduction in the motor time constant, (c) a sizeable decrease in time delay for one subject, and (d) an apparently large increase in motor noise for the other subject. Surprisingly, training had small and inconsistent effects on observation noise related to utilization of error rate information.

The cross-comparison significance test was applied to determine which of the identified parameter changes reflected real differences in operator behavior. Tests were performed for the following sets of parameters: (a)

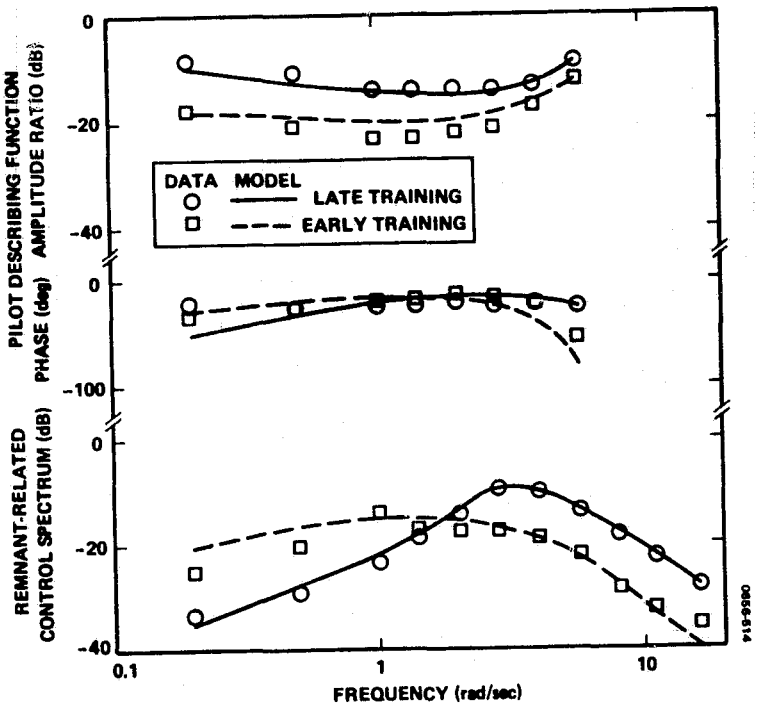


Figure 1. Effects of Training on Frequency Response
Subject CP, average of 4 trials.

the entire set, (b) observation noise parameters as a group, (c) motor time constant, and (d) time delay. Modeling error ratios were computed separately for subjects CP and TB.

Table 3 shows that, taken as a whole, changes in pilot-related model parameters were highly significant. Average model parameters yielded matching errors that were from about 8 to 20 times as great as those obtained with the optimal parameter sets. This result is not surprising, given the substantial training-related changes in operator response behavior shown in Figure 1.

Training-related differences in both the motor time constant parameter and the noise parameter group were important. Differences associated with motor time constant were more significant in the sense that error ratios for this grouping were about 50% higher than ratios associated with the noise parameters.

Fixing the time delay by itself yielded error ratios only slightly greater than unity. A test performed on the large training-related change

Table 3. Test of Model Parameter Differences Due to Training

	Modeling Error Ratio	
Parameter Set Tested	Subject CP	Subject TB
All Parameters	18.5	7.8
All Noises	3.2	2.0
Motor Time Constant	5.0	3.6
Time Delay	1.1	1.0

in motor noise found for subject TB also yielded negligible change in modeling error ratio. Thus, training-related effects on identified changes in motor time constant and observation noise appear to reflect true changes in operator response capabilities, whereas changes in time delay and motor noise are more likely to reflect (for these specific data sets) problems in parameter identification.

DISCUSSION

The requirement for a given parameter to be included in the identified set, and the ease and precision with which the parameter can be identified, are not intrinsic properties of the model parameter in question. Rather, these factors depend partly on the details of the task structure and of the analysis procedures. For example, we showed above that motor noise was required to obtain minimum matching error for one task but not for another.

Parameterization and identifiability will also depend strongly on the experimental measurement set used to define the matching error, and on the set of model parameters being identified. For example, sensitivity analysis performed in other studies [8] suggests that omission of the remnant spectrum from the measurement set would lead to considerable difficulty in distinguishing among the various observation noise sources (and possibly in distinguishing observation noise from time delay). Similarly, if one were to attempt to identify cost weightings for all state variables, along with the independent pilot parameters considered in this paper, overparameterization might well impede identification of one or more parameters.

Caution should be exercised when interpreting the training-related changes in model parameters reported above. As in previous studies, model analysis was based, in part, on the assumption that the subject has a near-perfect internal representation of the task environment (plant dynamics, input spectrum, etc.). While this assumption is appropriate for well-trained subjects tracking with relatively low-order plants, it is less likely to apply to subjects early in training.

More comprehensive analysis of the data base suggests that training-related changes in motor time constant do not reflect differences in motor response capabilities, but other kinds of response limitations not adequately reflected by the model as applied to this study(Levison, 1981). For example, the large motor time constant found early in training may reflect a cautious control strategy (i.e., low pilot gain) arising from the subject's uncertainty with regard to the dynamical response characteristics of the controlled element. Further research is contemplated in this area.

ACKNOWLEDGEMENT

This work was supported by the Air Force Office of Scientific Research under Contract No. F49620-8-C-0073.

REFERENCES

1. Levison, W.H., "Effects of Whole-Body Motion Simulation on Flight Skill Development", Report No. 4645, Bolt Beranek and Newman Inc., Cambridge, MA, 1981 (in preparation).

2. Baron, and W.H. Levison, "The Optimal Control Model: Status and Future Directions", Proceedings of the 1980 International Conference on Cybernetics and Society, Cambridge, MA, Oct. 8-10, 1980.

3. Lancraft, R.E. and D.L. Kleinman, "On the Identification of Parameters in the Optimal Control Model", Proceedings of the Fifteenth Annual Conference on Manual Control, Wright State University, Dayton, OH, March 1979.

4. Levison, W.H., "The Effects of Display Gain and Signal Bandwidth on Human Controller Remnant," AMRL-TR-70-93, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, OH, March 1971.

5. Levison, W.H., "Modeling the Pilot's Use of Motion Cues During Transient Aircraft Maneuvers", Report No. 4312, Bolt Beranek and Newman Inc., Cambridge, MA, March 1980.

6. Levison, W.H., R.E. Lancraft, and A.M. Junker, "The Effects of Simulator Delays on Performance and Learning in a Roll-Axis Tracking

Task", paper presented at the Fifteenth Annual Conference on Manual Control, Wright State University, Dayton, OH, March 1979.

7. Phatak, A., H. Weinert, I. Segall, and C. Day, "Identification of a Modified Optimal Control Model for the Human Operator", Automatica, Vol. 12, pp. 31-41, 1976.

8. Zacharias, G.L. and W.H. Levison, "A Performance Analyzer for Identifying Changes in Human Operator Tracking Strategies", ARML-TR-79-17, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, March 1979.

INFORMATION AND DISPLAY REQUIREMENTS FOR AIRCRAFT TERRAIN FOLLOWING†

by

D. L. Kleinman, University of Connecticut, Storrs, CT.
J. Korn, ALPHATECH, Inc., Burlington, MA

SUMMARY

The objective of this work is to apply and validate the display design procedure for manned-vehicle systems, as premulgated in Refs [1]-[2], to a particular scenario of interest to the Air Force. The scenario chosen is that of zero-visibility high-speed terrain following (V = 466 ft/sec, H = 200 ft) with an A10 aircraft. We consider the longitudal (linearized) dynamics in our analysis. The variations in (commanded path over) terrain $\pi(t)$ are modeled as a 3rd - order random process.

The display design methodology is based on the Optimal Control Model of pilot response, and employs this model in various ways in different phases of the design process. The overall methodology, as shown functionally in Figure 1, indicates that the design process is intended as a precursor to manned simulation. It provides a rank-ordering of candidate displays through a three-level process.

1. *Information Level:* Here the OCM is applied to determine the relative importance of each system state to closed-loop performance, once a performance criterion is specified. For the candidate task, the performance criterion includes RMS terrain height errors $e(t) = \pi(t) - h(t)$ and RMS vertical acceleration. The OCM analysis indicates that error rate $\dot{e}(t)$ and terrain height acceleration $\ddot{\pi}(t)$ are the two most important pieces of information for the control task.

In addition to the above analysis, the OCM is also used to determine the optimal combination of system states to be used as a display for closed-loop control. This is the flight-director design process described in Ref [2]. For the candidate task the flight director is a linear combination of vehicle states and terrain shaping states.

2. *Display Level:* At this level of analysis the information requirements are integrated to propose several different realistic display systems. Human operator display and information processing limitations are included at this level, such as observation noise, attentional allocation, indifference thresholds, etc. For each candidate design, performance vs. workload curves are generated using the OCM. In the present case, the high utility of terrain

†Work supported by AF Aerospace Medical Research Laboratory under contract F33615-80-C-0528.

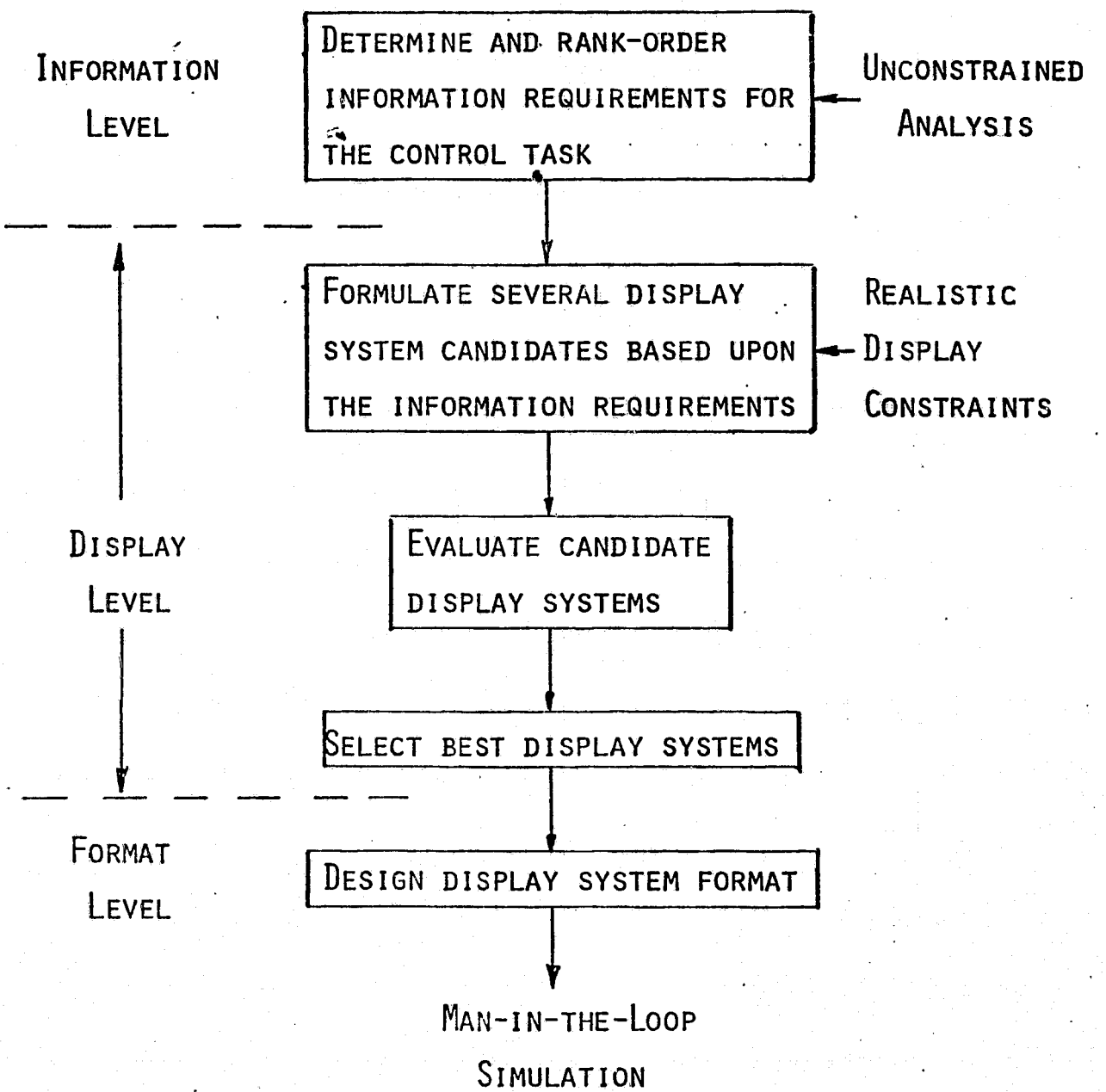


FIGURE 1 - OVERVIEW OF DISPLAY DESIGN METHODOLOGY

information (e.g. $\ddot{\pi}$) is included in a display via presentation/integration of the terrain height at a fixed distance/time ahead of the aircraft. Thus, displays are formulated to include realistically available measurements. Four candidate displays have been proposed via this procedure.

1. A tunnel display that indicates future desired flight path up to 3000' ahead of the aircraft.

2. A predictor-velocity vector display that indicates where the aircraft would be in $\simeq$ 4 sec, based on linear extrapolation, relative to the terrain height at that time.†

3. A combined tunnel plus velocity-vector display. This is basically the combination of 1) and 2).

4. The flight director display as designed at the information level.

In addition to the above synthetic display, the instrument panel is assumed to include: 1) Terrain height error e(t), 2) Radar altimeter, 3) pitch indicator and 4) G-meter.

The rank-ordering of the displays via the OCM indicates comparable levels of performance for displays 1-3, and much better performance using the flight director. Any of these displays yield significantly better performance than the non-synthetic display case, verifying the need for future flight path information.

3. *Format Level:* Here specific display formats are suggested for the presentation of specific display systems designed in level 2. Thus at this level the analytic display is translated into requirements for a physical display. Here we determine display layout, size and mode of presentation suitable for a man-in-the-loop simulation. The work at this level is largely an "art", but is guided by the sensitivities, attentional allocations, etc. that are generated by the OCM at the display level.

Man-in-the-loop experiments that evaluated the performance of the four candidate display systems were conducted at the CYBERLAB facility at the University of Connecticut. The experimental results tend to confirm the analytic rank-ordering of the candidate displays, and show a marked improvement in performance for the flight director design.

REFERENCES

1. Curry, R.E., D.L. Kleinman and W.C. Hoffman, "A Design Procedure for Control/Display Systems", Human Factors, Vol. 19, No. 5, Oct. 1977.

2. Hoffman, W.C., Kleinman, D.L. and Young, L.R., "Display/Control Requirements for Automated VTOL Aircraft", NASA CR-158905, Oct. 1976.

†This display is similar to the terrain box in use on the A10 HUD.

A DATA COLLECTION SCHEME FOR IDENTIFICATION OF PARAMETERS IN A DRIVER MODEL

By B. W. Mooring, M. McDermott, and Je-Meng Su

Texas A&M University

SUMMARY

When adapting a vehicle for use by a handicapped driver, it is often necessary to employ a high gain steering controller to compensate for limitations in the driver's range of motion. Because such a driver/vehicle system can become unstable as vehicle speed is increased, it is desirable to use a computer simulation of the driver/vehicle combination as a design tool to investigate the system response prior to construction of a controller and road testing. While there are a number of different driver models in existence, they all contain some unknown driver parameters which must be identified prior to use of the model for system analysis. This work addresses a means to collect the data necessary for identification of these driver model parameters without extensive instrumentation of a vehicle to measure and record vehicle states.

The procedure consists of three steps. First, a road test is conducted with the driver in a normal vehicle, during which only the steering wheel angle and the vehicle speed is recorded. Next, the data from the road test is input into a computer model of the vehicle which integrates the vehicle equations of motion with the given speed and steering inputs to yield the vehicle states, some of which the driver senses. Finally, with the sensed vehicle states as inputs and the recorded steering response as output, a least squares parameter identification procedure is used to compute the parameters in the proposed driver model.

Initial tests of the procedure identified all of the driver parameters with errors of 6% or less.

INTRODUCTION

The Rehabilitation Engineering Program at Texas A&M University is currently involved in the evaluation and design of low effort - high gain automotive control devices for handicapped drivers. Experience has shown that some vehicles with high gain steering are relatively easy to control at moderate speeds and others require maximum driver effort to maintain control at very low speeds. In order to better understand the causes of this wide variation in handling proterties and to quantify the effects of changes in various steering system parameters on vehicle response, development of a computer simulation of the driver-vehicle combination was begun.

At present, there are a number of vehicle models [3] and driver models [1,5,6] that are available. Use of a typical vehicle simulation requires knowledge of the geometry, inertial properties, and tire characteristics of the vehicle to be studied. Most of these can be obtained by direct measurement or are easily estimated.

As with the vehicle models, most of the driver models were found to have several coefficients whose magnitude is dependent on the characteristics of the driver or his environment. In order to use a driver model, a means to quickly and inexpensively identify these unknown parameters in a driver model is required. Classically, this identification is accomplished by running a road test and recording the driver inputs and all of the vehicle motion variables that the driver may respond to. As illustrated in Figure 1a, the vehicle motion data is used as an input for the driver model and the driver response data is compared to the results of the driver model to generate an error. There are a number of parameter identification techniques available [4] to determine the coefficients in the driver model that will minimize the error in predicted and measured driver response. The difficulty in employing this procedure lies in the instrumentation required to record the vehicle motion variables. Variables such as heading angle or yaw rate require a gyroscopic device which is expensive and bulky. The position of the vehicle on the road may be obtained with an optical tracking device or by integrating the output of accelerometers on an inertial platform. Either of these methods is expensive and the equipment is not easily moved from one vehicle to another.

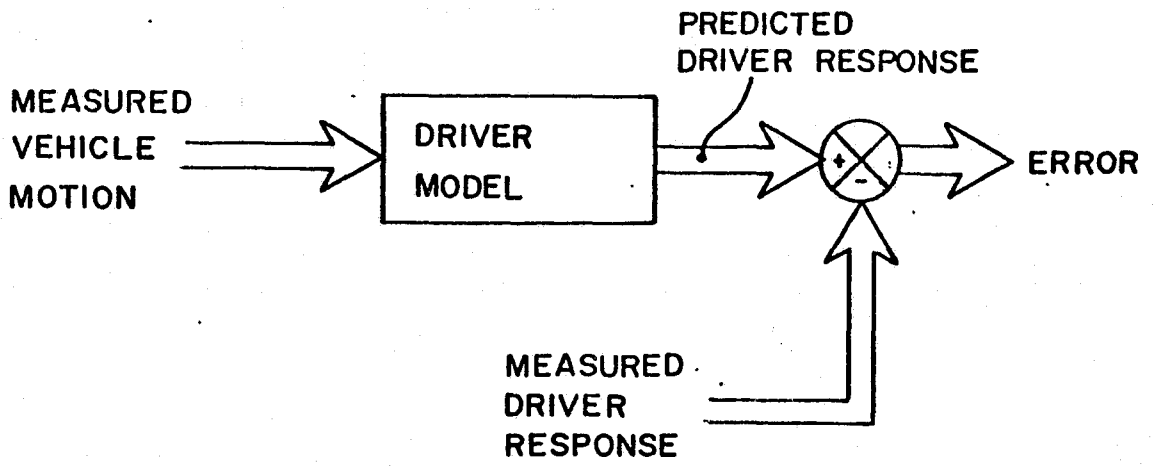


Figure 1a. Typical Parameter Identification Procedure

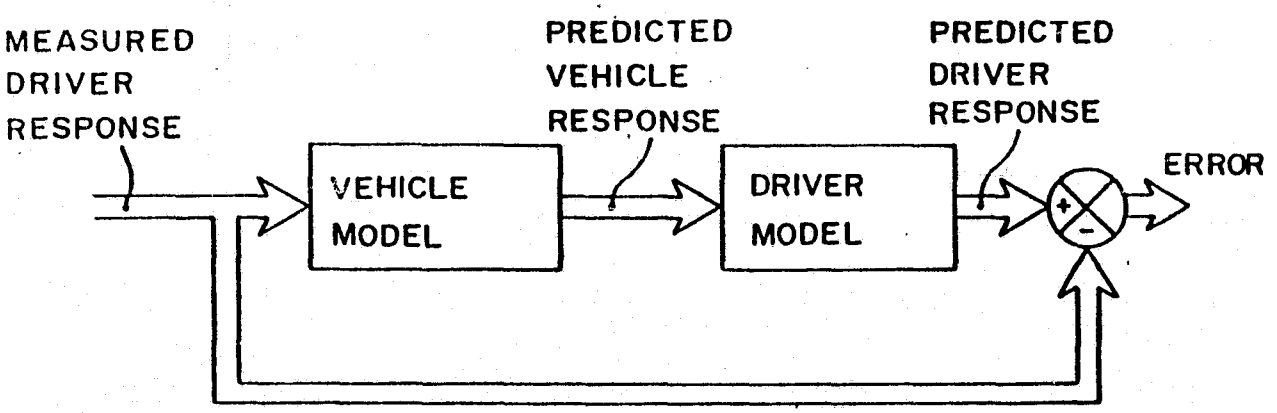


Figure 1b. Proposed Procedure

In order to minimize the instrumentation necessary to obtain the required data, an alternate procedure was examined. As illustrated in Figure 1b, the only measured variables are driver inputs (steering wheel angle and vehicle forward velocity). These inputs are then used with a vehicle simulation to determine what the motion must have been during the test. In this manner the required data for driver model identification is obtained with a minimum of instrumentation.

VEHICLE MODEL

The vehicle model used in this procedure was chosen because it was considered to be the simplest model available that exhibited the handling properties under study. Figure 2a is an illustration of the vehicle model. This vehicle has three degrees of freedom including the forward and lateral position of the mass center P, (r_{PX} and r_{PY}) and the heading angle (ψ). Two coordinate systems are employed. The X, Y, Z system is fixed to the roadway and has the associated unit vectors $\hat{I}$, $\hat{J}$, and $\hat{K}$. The x, y, z system is fixed to the car and has unit vectors $\hat{i}$, $\hat{j}$, and $\hat{k}$.

Using these definitions, the acceleration of the mass center, P, may be shown to be

$$\bar{a}_p = [\ddot{r}_{PX} \cos \psi + \ddot{r}_{PY} \sin \psi]\, \hat{i} + [\ddot{r}_{PY} \cos \psi - \ddot{r}_{PX} \sin \psi]\, \hat{j} \qquad (1)$$

The free body diagram for the vehicle is illustrated in Figure 2b. In order to simplify the analysis, secondary forces such as tire rolling resistance, self aligning torque, aerodynamic drag, and gyroscopic moment are considered negligible. Applying Newton's Laws to the free body diagram results in the following equations of motion.

$$F_{fL} \cos \delta_L + F_{fR} \cos \delta_R + F_{RL} + F_{RR} = m\ddot{r}_{PY} \cos \psi - m\ddot{r}_{PX} \sin \psi \qquad (2)$$

$$-a_5(F_{RL} + F_{RR}) + a_1(F_{fL} \cos \delta_L + F_{fR} \cos \delta_R) - (a_2 + a_4)\, F_{fL} \sin \delta_L$$

$$+ (a_2 - a_4)\, F_{fR} \sin \delta_R = I\ddot{\psi} \qquad (3)$$

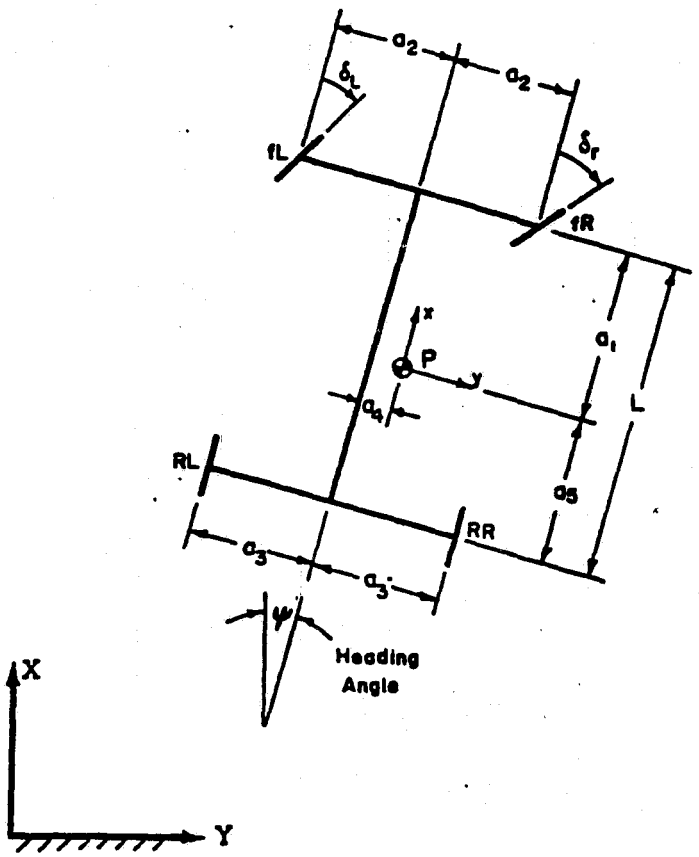


Figure 2a. Vehicle Model

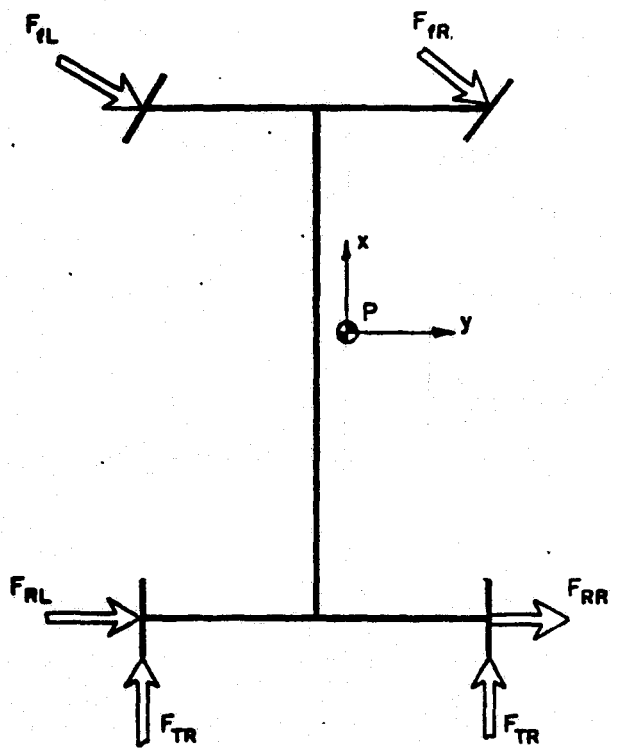


Figure 2b. Vehicle Free Body Diagram

A third equation may be obtained by recalling that the vehicle forward velocity is known. The x component of $\bar{a}_p$, therefore, may be expressed as

$$\ddot{r}_{PX} \cos \psi + \ddot{r}_{PY} \sin \psi = \dot{V} \tag{4}$$

Equations (2), (3), and (4) represent the three equations of motion for the vehicle. Unfortunately, the tire forces (F_{fL}, F_{fR}, F_{RL}, F_{RR}) are unknown and must be related to the vehicle motion parameters. It is well documented [2] that a rolling pneumatic tire under the influence of a lateral load exhibits a viscoelastic deformation of the tread surface. This deformation results in a deviation of the wheel center velocity from its expected direction. This deviation is termed the slip angle and may be expressed in terms of the vehicle motion variables. Using the defined coordinates, (r_{PX}, r_{PY}, and ψ) the slip angle at each wheel may be shown to be:

$$\alpha_{fL} = \delta_L - \tan^{-1}\left[\frac{\dot{r}_{PY} \cos \psi - \dot{r}_{PX} \sin \psi + a_1\dot{\psi}}{V + (a_2 + a_4)\dot{\psi}}\right] \tag{5}$$

$$\alpha_{fL} = \delta_R - \tan^{-1}\left[\frac{\dot{r}_{PY} \cos \psi - \dot{r}_{PX} \sin \psi + a_1\dot{\psi}}{V - (a_2 - a_4)\dot{\psi}}\right] \tag{6}$$

$$\alpha_{RL} = - \tan^{-1}\left[\frac{\dot{r}_{PY} \cos \psi - \dot{r}_{PX} \sin \psi - a_5\dot{\psi}}{V + (a_3 + a_4)\dot{\psi}}\right] \tag{7}$$

$$\alpha_{RR} = - \tan^{-1}\left[\frac{\dot{r}_{PY} \cos \psi - \dot{r}_{PX} \sin \psi - a_5\dot{\psi}}{V - (a_3 - a_4)\dot{\psi}}\right] \tag{8}$$

where δ_L and δ_R are left and right steering angles as defined in Figure 2a.

Once the slip angle at a given wheel is specified, the tire lateral force may be determined. For this work, the tire used is a Goodyear FR70-14. Figure 3 illustrates the relationship between slip angle and lateral force for various normal loads on the tread surface.

Given the equations of motion (2-4), the slip angle expressions (5-8) and the tire force function illustrated in Figure 3, the motion of the vehicle may be determined if the forward velocity V, and the front wheel angles, δ_L and δ_R are known. Assuming Ackermann Steering, δ_L and δ_R may be related to the steering wheel angle by the following relationships

$$x = \frac{L}{\tan (K_{ST}\delta)} \tag{9}$$

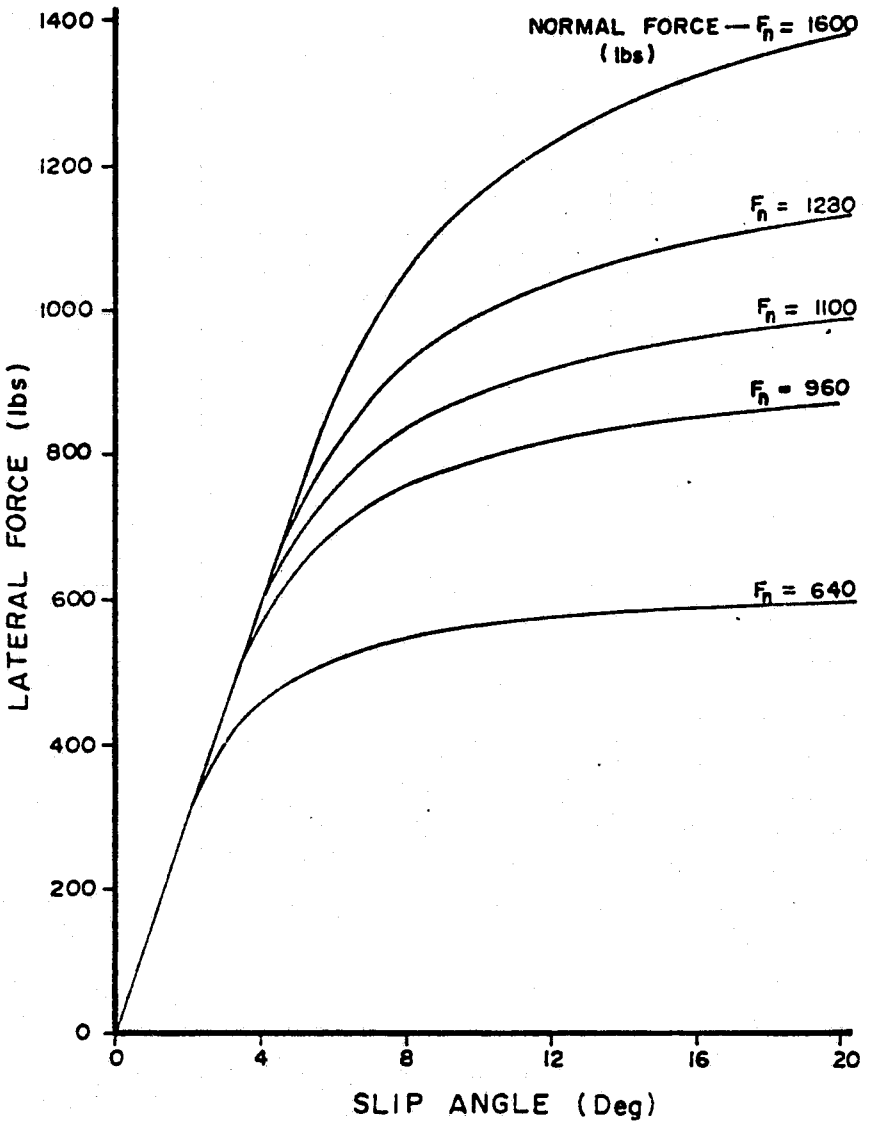


Figure 3. Tire Lateral Force vs. Slip Angle

$$\delta_R = \tan^{-1} \frac{L}{(x - a_2)} \tag{10}$$

$$\delta_L = \tan^{-1} \frac{L}{(x - a_2)} \tag{11}$$

where K_{ST} represents the steering gain, δ is steering wheel angle, and L is wheelbase.

Now, given the driver inputs of steering wheel angle, δ, and forward velocity, V, the motion of the vehicle may be determined. Because of the non-linearity of the motion and constraint equations, the most expedient means of obtaining a solution is by utilizing a numerical integration procedure on a digital computer. The computer that was used in the following example was a PRIME 750 minicomputer. The program was written in FORTRAN and utilized a predictor-corrector numerical integration scheme.

VALIDATION OF VEHICLE MODEL

In order to verify the accuracy of the vehicle simulation, a series of test runs was made with a vehicle that was instrumented to record several of the vehicle motion parameters as well as the driver responses. The driver responses were used as input data for the vehicle simulation. The results of the vehicle simulation were then compared to the vehicle motions recorded during the test. Several test runs were made at different vehicle speeds and over different courses. Figure 4 illustrates the comparison between predicted and measured heading angle, yaw rate, and lateral acceleration as the vehicle passed through an offset alley. As shown in Figure 4, the vehicle simulation provides a good estimation of the vehicle motion. The large difference in measured and predicted heading angle is due to a difference in reference position. The simulation automatically sets the vehicle's initial heading to zero degrees while the measured heading depends on the vehicle orientation when the gyroscope unit is switched on. When this bias is removed, the results compare quite well.

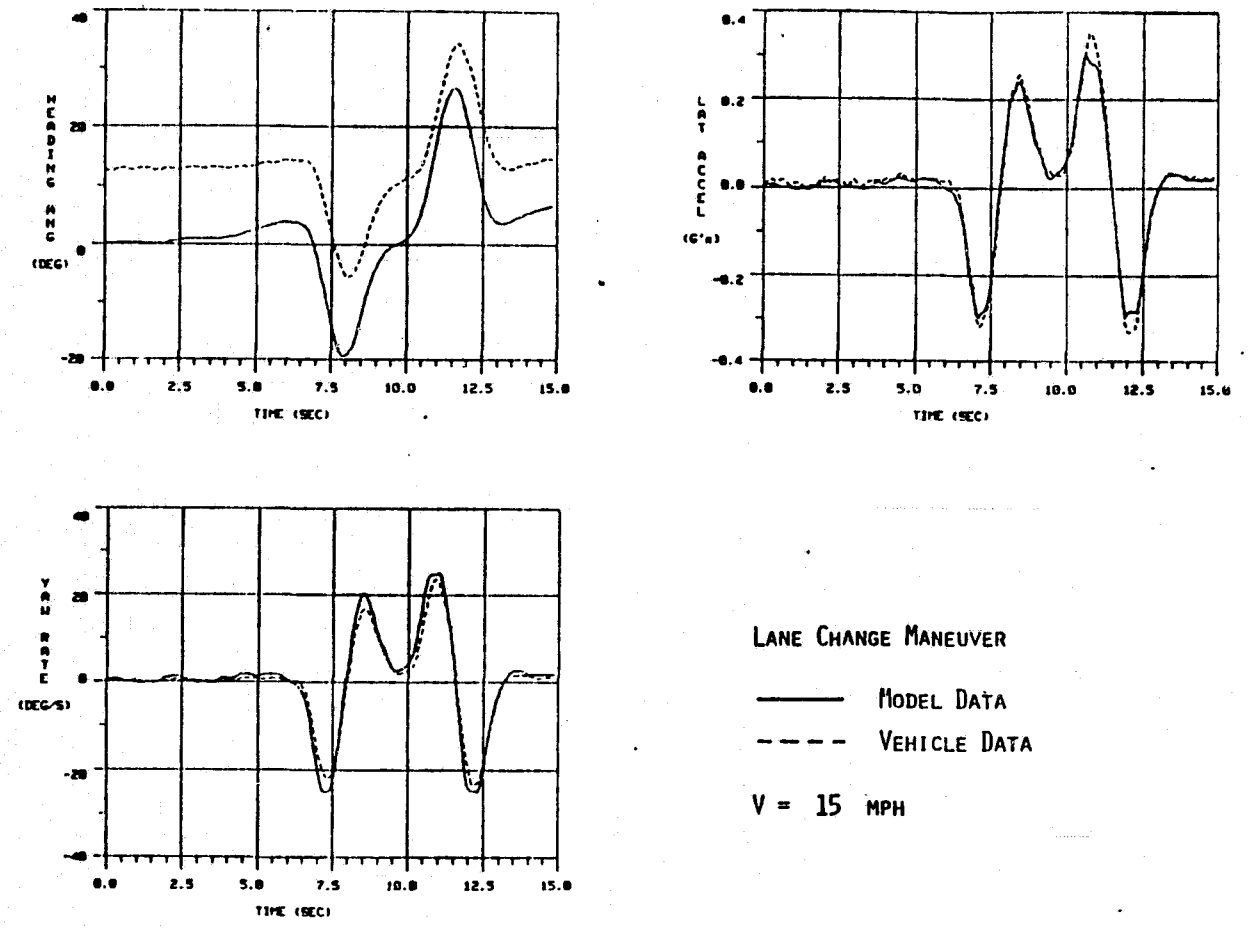


Figure 4. Comparison of Measured and Simulated Vehicle Motion

VALIDATION OF IDENTIFICATION PROCEDURE

The driver model used to validate the parameter identification procedure assumes that the driver can be represented as a two part cascade of a brain response and a neuromuscular lag. Reaction time delay and precognition or preview are not included since they tend to cancel and, at any rate, the particular form of the driver model is not critical to the identification procedure at this stage. The inputs to the driver model are the lane position error, ε_y, and the handing angle error, ε_ψ. For a straight roadway $\varepsilon_y = r_{py}$ and $\varepsilon_\psi = \psi$. The output of the brain is the commanded steering wheel angle δ' and is modeled as

$$\delta' = K\,(r_{py} + \eta\psi) \tag{12}$$

where K is the brain gain and η is a weighting factor. The brain output is the input to the neuromuscular system which is modeled as

$$\dot{\delta} = \frac{1}{\tau}\,(\delta' - \delta) \tag{13}$$

where τ is the neuromuscular time constant. Thus, the parameter identification procedure must compute values of K, η, and τ for which the computer model output best fits the measured data in a least squares sense.

Since the driver parameters K, η, and τ are not available from an actual driver/vehicle test a computer simulation was used to generate test data that allows direct comparison of computed parameter values to true parameter values. The test data was generated using the same driver model as the identification code, but a different vehicle model. The vehicle model used was more complete and had been thoroughly checked out.

The test maneuver was lane tracking on a straight road at a nominal speed of 55 mph with an initial lane error of five feet. As shown in Table 1 the largest error between the true and computed parameter values is 6%.

Parameter	True Value	Identified Value	Percent Error
τ	0.1	0.100	0%
K	-0.1	-0.094	-6%
η	100	94.40	-5.6%

Table 1 Errors in Identified Values of the Driver Model Parameters

CONCLUSIONS

As shown in Figure 4, the computer model provides a very accurate simulation of the actual vehicle states. This is a critical result, since the data collection scheme proposed in this paper is based on the premise that vehicle states, which are inputs to the driver, can be accurately constructed from simulation, thus eliminating the need to instrument each vehicle to be tested with an inertial platform. Also, assuming that an appropriate driver model is used, the identification procedure accurately computes the driver model parameters as indicated by the data in Table 1. Thus, each part of the overall procedure has been independently validated.

The final test of the procedure will be identification of driver model parameters from an actual road test and comparison of computed vehicle states and driver outputs to actual measured values. This testing is currently in progress.

REFERENCES

1. Baxter, J., Harrison, J. Y., "A Nonlinear Model Describing Driver Behavior on Straight Roads," Human Factors, 1979, 21(1), 87-97.

2. Clark, S. K., Mechanics of Pneumatic Tires, MBS Monograph 122, U.S. Government Printing Office, November 1971.

3. Ellis, J. R., Vehicle Dynamics, Business Books Ltd., London, 1973.

4. Graupe, D., Identification of Systems, New York, Robert E. Krieger Publishing Company, 1976.

5. McRuer, D. T., Allen, R. W., Weir, D. H., Klein, R. H., "New Results in Driver Steering Control Models," Human Factors, 1977, 19(4), 381-397.

6. Phatak, A. V., "Formulation and Validation of Optimal Control Theoretic Models of the Human Operator," Machine Systems Review, 1976, 2, 11-12.

2-D Results On Human Operator Perception

Alex A. Siapkara
Control Division
Hellenic Technology Center
402 Mesogeion Av.
Athens, Greece

Thomas B. Sheridan
Man-Machine Laboratory
Massachusetts Institute of Technology
Cambridge, Mass. 02139

ABSTRACT

This paper presents some preliminary results of the application of multi-dimensional scaling methodology in human factors engineering. The non-orthogonality of internally perceived task variables is exhibited for first and second order plants with both dependent and independent task variables. Directions of operator preference are shown for actual performance, pilot opinion rating, and subjective measures of fatigue, adaptability and system recognition. Improvement of performance in second order systems is, in addition, exhibited by the use of bang-bang feedback information. New dissimilarity measures for system comparison are suggested in order to account for human operator rotations and subjective sense of time.

1. INTRODUCTION

In comparing the objective performance of a human operator (H.O) with his/her subjective evaluations it was found helpful in reference [1] to use the methodology provided by multi-dimensional scaling (MDS). According to this methodology, first proposed in detail by Torgerson in 1952 [2] in the area of mathematical psychology, unidimensional pilot opinion rating scales (POR) are only to be thought of as the vecotrs in a multi-dimensional space which represents the "perceptual model" the H.O. has "internally constructed" on the nature and performance of his/her own task.

This internal task space (ITS) is conceptually different from the ones arrived at by deviate internal models of the H.O., as for example done in reference [3], in that the MDS formulation does not rely on the LQG models of the H.O. and is in essence completely model-free. What this ITS really depends on is the metric assumed to apply in constructing it from comparison measurements provided subjectively by the H.O. and interpreted mathematically via the MDS formulation. These comparisons fix the distances between the compared objects within the multi-dimensional space and correlate with the unidimensional subjective (SE) and objective evaluations (OE) of various system parameters via the use of a vector-fitting algorithm.

A complete analysis of the mathematical model employed including the real-time computational aspects involved in updating the ITS can be found in reference [4a]. The dimensionality of the matrix POR depends on the total number of tasks, compared, and more precisely

$$R^n \ni n \geq N \ \& \ \log_N (n_{max} + 1) = \bar{n}$$

where n the dimensionality of the ITS
N the number of task variables
n_i the # of selected values for the i^{th} variable, and
$\bar{n}$ the geometric average of n_i's.

This logarithmic complexity of MDS experiments makes it imperative to:

i. partition the original set of tasks into partially overlapping smaller groups which will still retain the property of clustering them back together "under the same roof" of a single ITS;
ii. design separate experiments with the smaller groups employing a small number of selected values for each variable;
iii. reduce mathematically via a statistical hyperplane fit the total number of dimensions towards the number of actual task variables (if these are known beforehand and the experiments well controlled).

The prescribed technique has been successfully utilized by Siapkara [4b] in performing a series of three sets of experiments with three subjects and two task variables. The resulting two-dimensional ITS, though certainly of limited validity, does effectively portray many of the associated MDS issues. Section 2 of this paper describes the experiments and section 3 deals with the experimental results. All accompanying figures can be found at the end of the paper.

2. THE SCOPE OF THE EXPERIMENT

Figure 1 shows in a self-explanatory way the experimental set-up used. The signal source is a smoothed slowly-varying output from a random number generator (RNG), so that the configuration of the systems to be controlled [Figure 2] corresponds to both a first and second-order system. Which of the systems is actually controlled by the H.O. depends on which of the two displays is accessible for consultation, and not on any difference in plant dynamics. In order to keep a small number of variables at hand, the parameters of both systems do not vary independently [Figure 3]; they are all instead in terms of a common parameter λ.

The second task variable K comes from the input. An effort was made to use a random-number generator which experiences minimal statistical variations over time and different starting values, so that only its mean strength (amplitude) will count. The generation method was based on a combination of techniques proposed by references [5] and [6]. Consistency of the smoothed random input was checked for two completely

different sets of task variables and measurement policies. [Figure 4]. It also checked comparatively well with the sinusoid method used in reference [7], with its statistics attaining an asymptotic stability much sooner, though the variations of the sinusoidal RNG are less frequent.

The use of fixed analog displays did not allow for the appropriate scaling of feedback indication, as would be the case with a flexible CRT screen. So, instead of performing the image range calibrations as outlined by equation (3) of reference [1], it was necessary to keep the number of display overshoots as an objective performance index in addition to tracking error scores. The spring characteristics of the vertical level knob used in the experiment was what Dommasch [8] would call a "bungee" (or down spring) control element, with a built-in center-wards pull which requires a constant off-center push in part of the subject in the case of proper control-force static stability. Transformation of the results to a situation which uses different control-element characteristics can be done via Rothbauer [9] [Figure 5]. Performance curves of these types of controlled elements is given in reference [10] [Figure 6].

Finally, provision was made in the experimental set-up to include a white-red light depending upon whether the tracking error was positive or negative. Experiments based solely on this type of feedback information will here-in be referred to as sign experiments. Justification for the inclusion of such an experiment is based on previously acquired insight in the field of experimental psychology. Johnson [11] mentions a 28-person 1968 experiment where multi-dimensional judgements correlated well with uni-dimensional equivalents when simple combinatory transformations of the various variables were performed. It was found that 42% judged on linear scales, 10.5% on quadratic, 45.5% on signed cues, and the rest in other configural modes. The signed cues were indicators with either +1 or -1 values, and they contained the sign information of the judgement only. For example, in the MDS context of paired comparisons the subjects would actually judge as if the stimuli were near the vicinity of just-discriminable differences: instead of rendering refined estimates of similarity, they would rather ask themselves some more foundamental questions: "Are the two situations compared different enough so as to bother giving out discrete and even more so continous estimates on scales beyond binary? And, if I admit there is a perceptible difference between them, will I be able later on on subsequent pair comparisons to maintain some credible consistency of how I rate these minute differences? Or, is it possible that I am going to develop the tendency of accentuating the dissimilarity scale near the similarity end of it, and compress thus my judgements on the truly dissimilar cases?"

3. EXPERIMENTAL RESULTS

Nine first and nine second order systems in all were evaluated. They were gotten by combining three values from each of the two variables

as shown in Figure 7. The same figure shows also the uncertainties and just-discriminable differences related with each of these task variables. Various types of subjective judgements rated on a 1-10 scale were collected in the experiments for these systems. The verbal characterization of these scales was only fixed at the ends, and is presented in Figure 8. For each category of systems there were essentially three kinds of runs performed: combined familiarization-evaluation runs, dissimilarity runs, and finally identification runs. For each run, in addition to the particular subjective judgement aimed at, objective performance records were kept, as well as subjective judgements on the level of fatigue experienced with the experiment and the level of difficulty (effort) in pronouncing the subjective evaluations themselves.

F a m i l i a r i z a t i o n - e v a l u a t i o n r u n s lasted 2 mins for each system. In the first 20 secs of familiarization SE on the success of familiarization was taken verbally around every 5 secs [Figure 9, located under Figure 6]. For 1-st order systems as $\lambda\uparrow$ so in general did the difficulty for familiarization (ex. VIII), but in no circumstance did the degree of familiarization decrease as time progressed. On the contrary 2-nd order systems with $K\downarrow$ experienced such a drop (ex. VII). The next 10 secs were considered a break between the familiarization portion of the run and the evaluation portion. During this interval a combined SE was provoked that indicated the efficacy of this relaxation period and the degree of comfort the subject experienced [Figure 10(a) ——]. The direction of relaxation increase is definitely different for 1-st and 2-nd order systems. The middle 60 secs of the runs are devoted to system evaluation. The subject with uninterrupted concentration on the task was previously instructed to perform his best. At the end of this period he/she gives an overall POR on the task [Figure 10(b) ---] , while OE are recorded for future comparison [Figure 10(b)——]. While OE increases with $K\downarrow,\lambda\downarrow$ for both 1-st and 2-nd order systems, SE's in general do not. Separated areas indicate entries which did not comform with the general direction of the property vectors and differ by more than one point in the psychological scale from what would have been considered as a value in acceptable deviation from the rule.

Along with the POR, the subject indicates his/her own mental and dexterity fatigue status, and the effort expended by him/her in rendering these SE's [Figure 10(c)]. A general kind of agreement can be seen for both SE measures used, while the scales utilized by the subject differ by one and two points, the subject being more harsh on rating the fatigue factor. The last 30 secs of these runs are used for deadaptation purposes till the start of the subsequent run. Figure 10a (vectors in segmented line) shows the degree of comfort felt by the H.O. It can be seen that inter-run deadaptation does not relate to intra-run relaxation, though both are comfort accomodating. This is so because relaxation on the same task is viewed by the H.O. simply as a means to reduce his/her fatigue, whilst deadaptation seems to depend more upon the anxiety of what comes next.

D i s s i m i l a r i t y r u n s lasted 70 secs for each pair of systems compared. Each member of the pair was controlled for 30 secs at the end of which OE was recorded. Figure 11 shows the variations in performance relative to the OE of evaluation runs as a standard of reference for both types of systems. 10 secs in between the individual presentations served as a relaxation during which the degree of ease for remembering the

behaviour of the system presented first was recorded [Figure 10(d)]. The incomplete nature of the memorization space results from the factorial design of the dissimilarity runs which forms a minimum number of combinations to be compared. Despite the lack of additional information the essential character of the memorization vectors is evident, and suggests that the impression of remembering a system remains in direction the same for 1-st and 2-nd order systems. However, this does not necessarily mean that actual ability for such memorization is so. This point is discussed in more detail later on. At the end of the dissimilarity runs the H.O. judges the similarity of the systems compared, as well as his/her own degree of comfort in pronouncing this judgement. An off-line procedure for computing two-dimensional ITS's is invoked at the end of all the dissimilarity experiments.

The three sets of experiments involved tasks V_1-VI_1-IX_1, I_1-II_1-III_1-IV_1-VII_1-IX_1 and I_2-II_2-III_2-IV_2. After finding the task vectors based on these partitioned experiments, the ITS's are brought together by modifying the point dispersions so that the task vectors coicide [Figure 13(a)]. The rest of the systems $VIII_1$ and V_2-IX_2 are then placed within this combined ITS by a straight-forward two-dimensional interpolation procedure. The same is done with objective performance and subjective evaluation vectors for both 1-st and 2-nd order systems [Figures 13(b)+(c)].

4. DISCUSSION

Many could be the implications of these configurations, if it was not for the limited evidence for these internal space constructions. What is for sure is that the evidence collected on the few subjects of the experiments exhibits on the average tendencies which were sort of anticipated, and which motivated this study in the first place, anyway. These tendencies are:

i. the task vectors are not in general perceived independently,
ii. objective and subjective ratings of manual tracking tasks do not necessarily coincide.

The large inconsistencies in comparison distances (offsets that are 30% off from the Euclidean point of view), and the quite considerable variations in performance (uncertainties of the order of 36.4%) raised a number of questions on the validity/utility of MDS in the multi-dimensional assessment of POR's.

This motivated a third stage in the experiments beyond familiarization-evaluation and dissimilarity runs. More specifically, i d e n t i f i c a t i o n r u n s were specially conducted to test the hypothesis inherent in these experiments, that the H.O. could identify successfully the differentiating character of the systems he/she is confronted with. Figure 10d (vectors in segmented line) shows the directions of maximum increase in actual memorization, assuming that identifiability is stricty speaking a measure of memorization. In certain cases an almost complete lack of ability to identify a specific task is evident, as for example: III_1,VII_2,IV_2 and III_2. [Figure 12]. The short duration of the tasks can be cited as a main cause, because identifiability is a cumulative property which combines together the temporal elements of a task; and if the task has not fully developed its

essential idiosyncracies this plays a negative role on the H.O. perceiving its global nature. Figure 14 shows, for example, the immense variability in the character of an almost marginal 2-nd order task in 10 secs intervals. This particular task reveals its true collapsing character after 50 secs, and certainly not within the first 30 secs from its activation. On the other hand, difficulties in memorization are unrelated to the factorial design of the similarity experiments. Figure 15 clearly shows that fatigue accumulated on entire groups of comparisons (based on the same first member of the pairs) does not seem to relate directly to the difficulty associated with identifying those systems.

Finally, the improved performance shown with sign rather than complete information on tracking error [Figure 16(a)] motivated a correlation between the various systems and the number of bang-bang pieces of feedback information. Figure 16(b) suggests that the implicit strategy used by the H.O. in optimizing manual tracking performance is to try to reach a uniform level of acceptance in the number of tracking error crossovers.

5. FUTURE DIRECTIONS

Despite some of the disheartening aspects of applying MDS on POR pronouncement, this has much more to do with H.O. inconsistencies than with a methodological difficulty and/or inability inherent in MDS itself. To the contrary, reference [4c] suggests that an MDS approach to the problems of mental workload in a multi-task environment and of multi-operator judgement and control (collective task-attending) would simplify their study by avoiding the use of the law for comparative judgement [12] and of group probability partitioning [13] , respectively.

Relating individual ITS's of various tasks or various operators, —firstly between themselves and secondly with the more complex ITS's resulting from multi-task or multi-operator situations—, would in the opinion of the authors provide us with usuful POR matrix transformations; for example:

i. between groups of people with different levels of aptitude in performing manual tracking tasks,
ii. for increasing the reliability of the operation by providing feedback information to the H.O. about discrepancies in his/her ITS between objective performance and its subjective evaluation,
iii. for the design of flexible control/display configurations which will automatically adapt their dynamics according to the ITS peculiarities of the H.O. involved in the operation so as to improve performance in a way transparent to the H.O., etc.

Finally, it has to be noted that Euclidean or even Minkowski spaces provide metrics for ITS that are not suitable to express the contribution of terms that correspond to situations where H.O. space rotations and his/her subjective sense of time (the time thought of as having been ellapsed in the controlling action) play an important role. This is so because the base vectors representing a space rotation combine in a multiplication group [14] and those representing time belong to the spinor class [15]. This can be illustrated as follows:

Consider $\underline{y} = x_1\hat{e}^1 + x_2\hat{e}^2 + \ldots\ldots + x_n\hat{e}^n$

and $$\| y \| = [s_1 x_1^p + s^2 x_2^p + \ldots\ldots + s^n x_n^p]^{1/p} \qquad (1)$$

where only superscript p is a power. Then for variables x_i that constitute

i. an inner product group, $s^i = +1$

ii. elements of a space rotation,

$$\hat{e}^{k_1} \odot \hat{e}^{k_2} \odot \ldots\ldots \hat{e}^{k_{p-1}} = \hat{e}^{k_p}$$

where $k_1, k_2, \ldots\ldots k_p$ in circular order

iii. a designant for the H.O.$_i$ sense of ellapsed time (when experiments seem to depend on it), $s^i = -1$.

It can be seen that (1) applies only in the first and third cases, whereas the metric corresponding to the second case is given by the more general form

$$\| y \| = [\hat{e}^{k_1} \odot \hat{e}^{k_2} \odot \ldots\ldots \hat{e}^{k_p} \cdot x_{k_1} x_{k_2} \ldots\ldots x_{k_p}]^{1/p} \qquad (2)$$

following reference [14], the summation convention and the definition

$$\| y \| = [\underline{y} \odot \underline{y} \odot \ldots\ldots \odot \underline{y}]^{1/p}$$

The double-circle operator might be any legitimate operator, such as an inner or outer product, an integration of base functions, or even a convolution integral in the case of cascaded moving vectors, or a meaningful mixture of the above.

Whereas solving MDS with metrics of the form (1) seems to be a trivial extension of the case where $s^i = 1 \ \forall\, i$, this is certainly not the case with the much more general form (2), where coupled terms do in general appear.

ACKNOWLEDGEMENT

The authors would like to pay tribute to Dr. R.E. Curry of NASA Ames for shaping while at MIT in 1975 the original idea of using MDS in multidimensional POR, and suspecting the non-ortogonality of internally perveived task variables.

A. Siapkaras would, in addition, like to thank Dr. J.A. Argyris, Director of the ISD Aerospace Institute of Stuttgart University, FRG for partly releiving him from his contractual arrangements for preparation of the manuscript.

REFERENCES

1. Siapkara, A.A. "Multi-Attribute Subjective Evaluations of Manual Tracking Tasks vs. Objective Performance of the Human Operator" 13-th Annual Conference on Manual Control, June 1977, 118-125

2. Torgerson, W.S. "MDS Theory and Methods" Psychometrika 17 (1952), 401-419

3. Baron, S. and Berliner, J.E. "The Effects of Deviate Internal Representations in the Optimal Model of the Human Operator" Procs. of the 1976 IEEE Conf. on Decision & Control, 1055-1057

4. Siapkara, A.A. "Multi-Attribute Evaluations of Piloting Tasks" M.I.T. Eng.D Thesis, December 1978; (partially referred to as a = ps. 180-232, b = ps. 233-261, c = 262-287)

5. Zelen, Marvin "Random Number Generation" Handbook of Mathematical Functions, US/NBS 1970, 949-953

6. Hull + Dobell "Random Number Generators" SIAM 4 (1962), 230-254

7. Shirley, R.S. "A Comparison of Techniques for Measuring Human Operator Frequency Response" 6-th Annual Conference on Manual Control, 1970, 803-825

8. Dommasch "Airplane Aerodynamics" Ed. Pitman, GB 1961, 456

9. Rothbauer, Gunter "Influences of Joy-Stick Spring Resistance on the Execution of Simple and Complex Positioning Movements" 13-th Annual Conference on Manual Control, June 1977, 447-451

10. Frost, G. "Man-Machine Dynamics" The Van Cott & Kinkade U.S. Armed Forces Guide to Human Engineering Design, Ed. 1972, 236-241

11. Johnson, D.M. "Multi-Dimensional Judgements on Continuous and Discrete Scales" Systematic Introduction to the Psychology of Thinking Ch. 8, p. 379, Harper & Row 1972

12. Daryanian, B. "Subjective Scaling of Mental Workload in a Multi-Task Environment" 16-th Annual Conference on Manual Control, May 1980, 172-188

13. Ferrell, W.R. "A Model of Subjective Probabilities from Small Groups" 16-th Annual Conference on Manual Control, May 1980, 271-284

14. Glazman, I.M.; Ljubic, J.I. "Finite-Dimensional Linear Analysis", The M.I.T. Press 1974, (§5.6, §6.3)

15. Cartan, Elie "The Theory of Spinors" The M.I.T. Press, Mono. 37, 1966 (§5.6, §6.6)

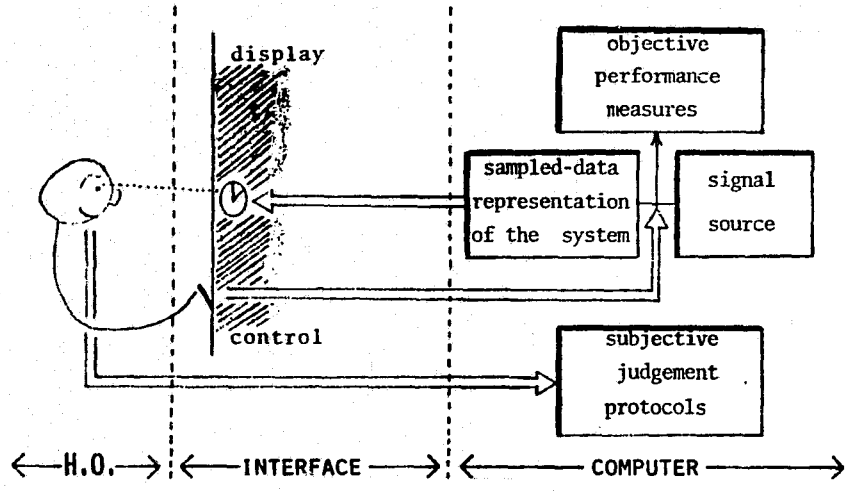


FIGURE 1

THE EXPERIMENTAL SET-UP

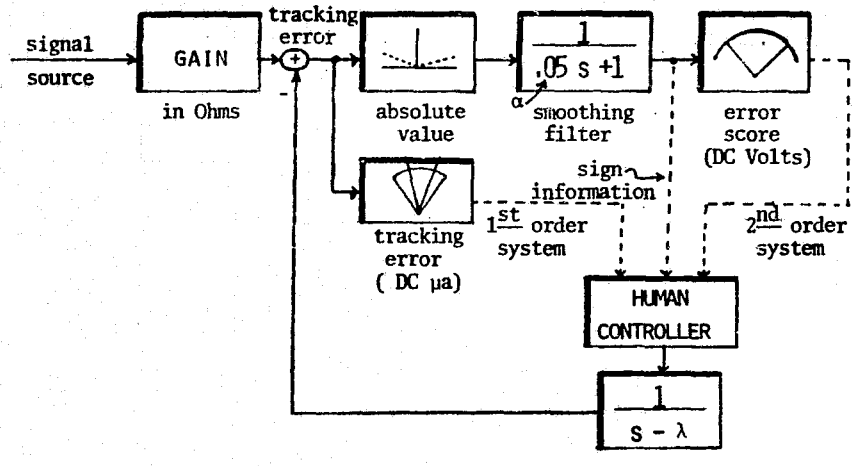


FIGURE 2

THE SYSTEM TO BE CONTROLLED

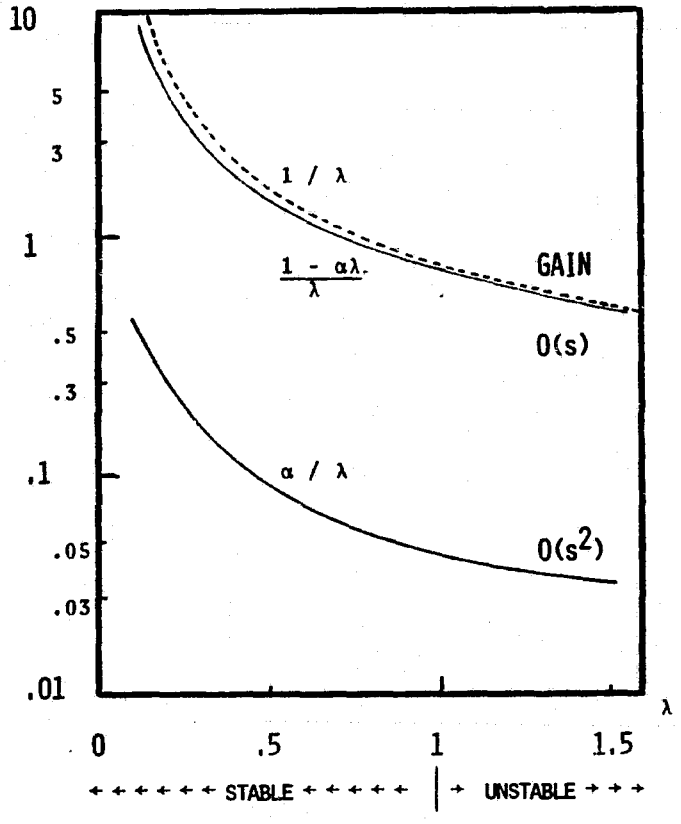


FIGURE 3

THE VARIABILITY OF SYSTEM PARAMETERS

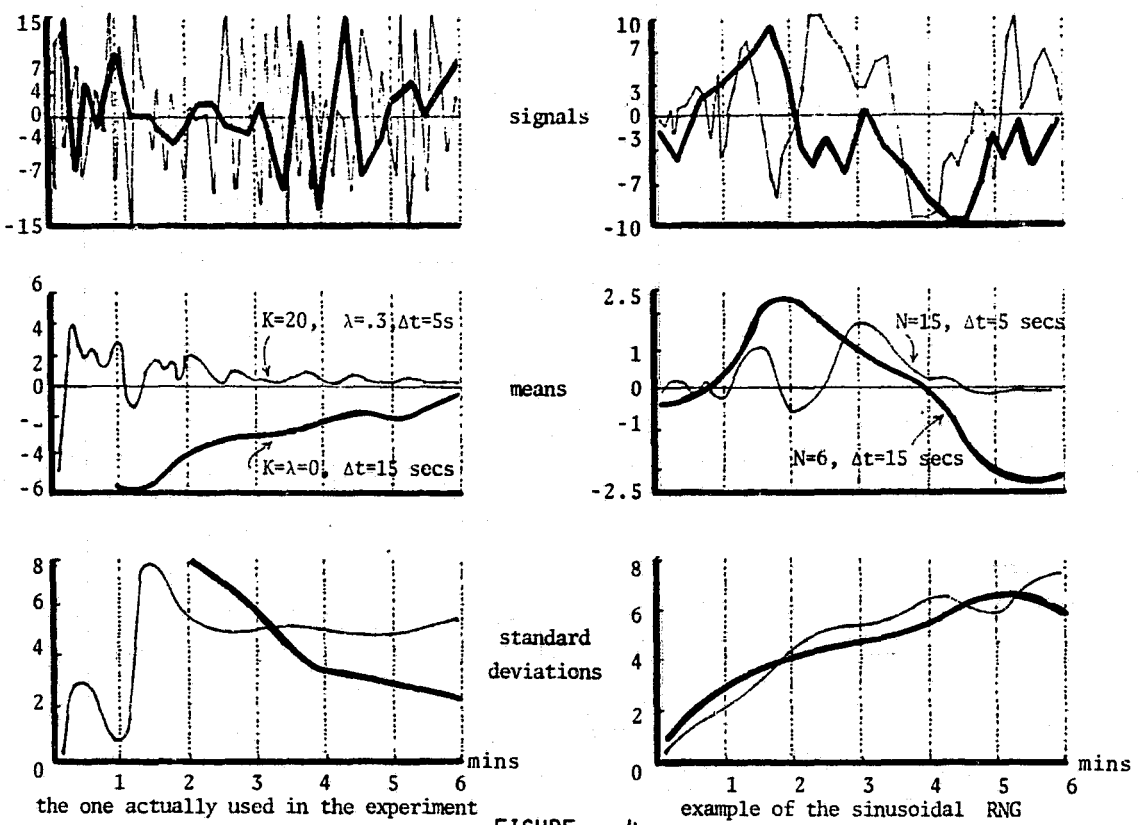


FIGURE 4

CHARACTERISTICS OF THE SMOOTHED RNG

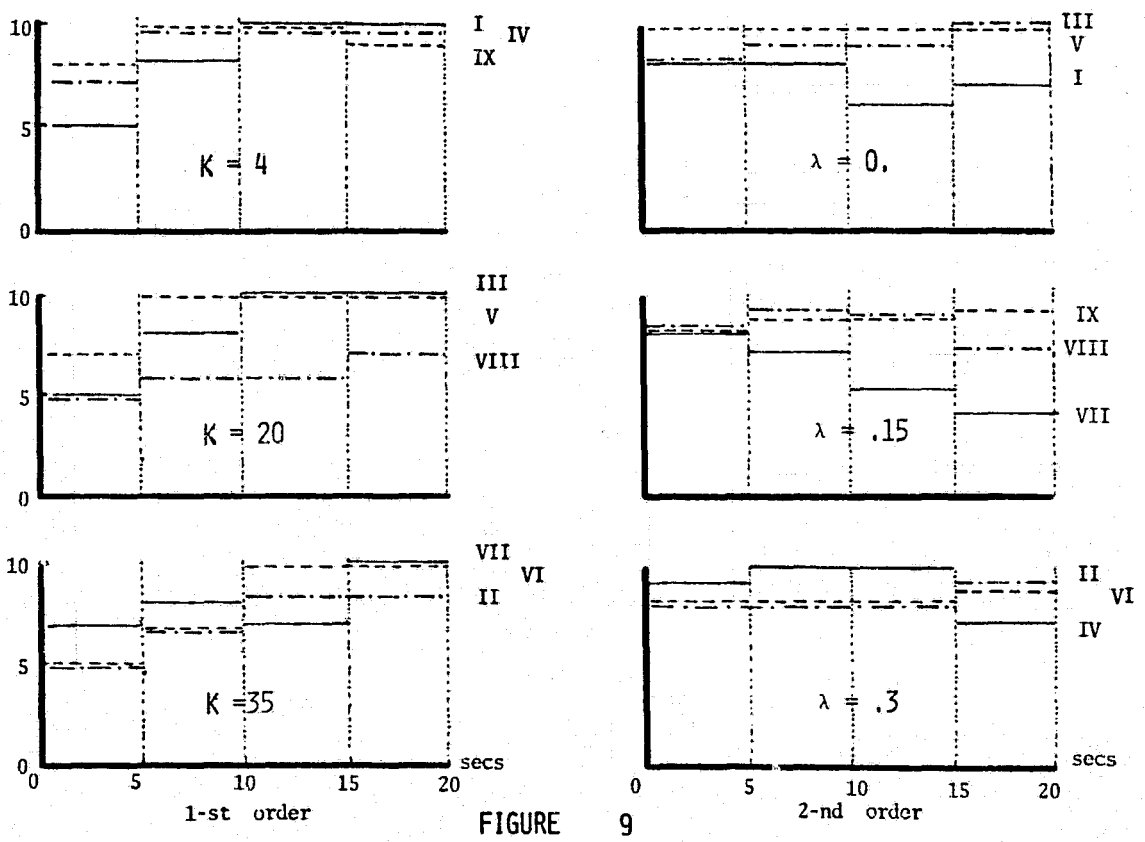


FIGURE 9

HISTORY OF SUBJECTIVE EVALUATIONS ON FAMILIARIZATION

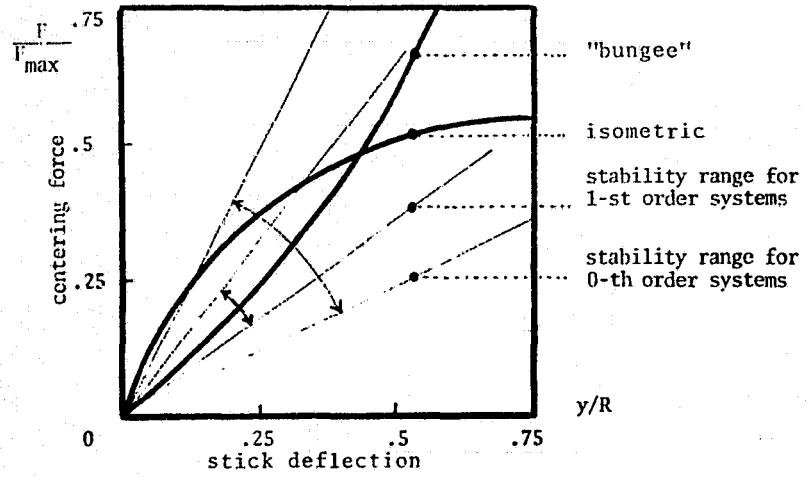
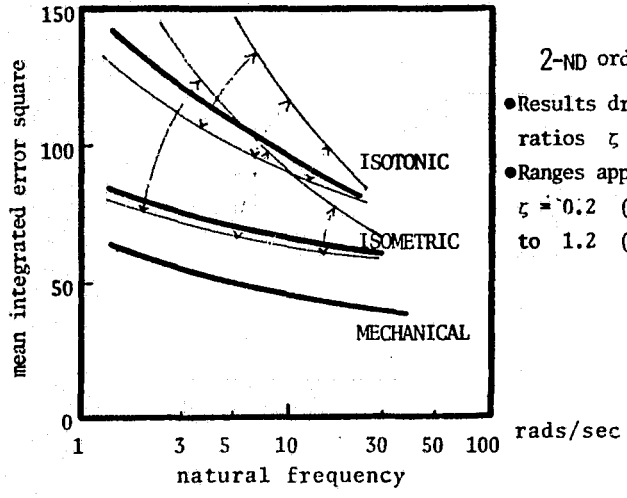


FIGURE 5

CONTROLLED ELEMENT TRANSFORMATIONS
(adopted from ROTHBAUER)



2-ND order systems

- Results drawn for damping ratios $\zeta = .8$
- Ranges apply for the region $\zeta = 0.2$ (upper margin) to 1.2 (lower ")

FIGURE 6

PERFORMANCE CURVES FOR VARIOUS CONTROLLED ELEMENTS

1st order

λ \ K	4	20	35
.1	I	III	VII
.3	IX	V	VI
.5	IV	VIII	II

2nd order

λ \ K	0	10	20
0.	I	V	III
.15	VII	IX	VIII
.3	IV	VI	II

thresholds

K	2	5
λ	.025	.1
•	±2secs	5for ∀ t : t≈30

time intervals

jdd in perceiving

uncertainty in the settings

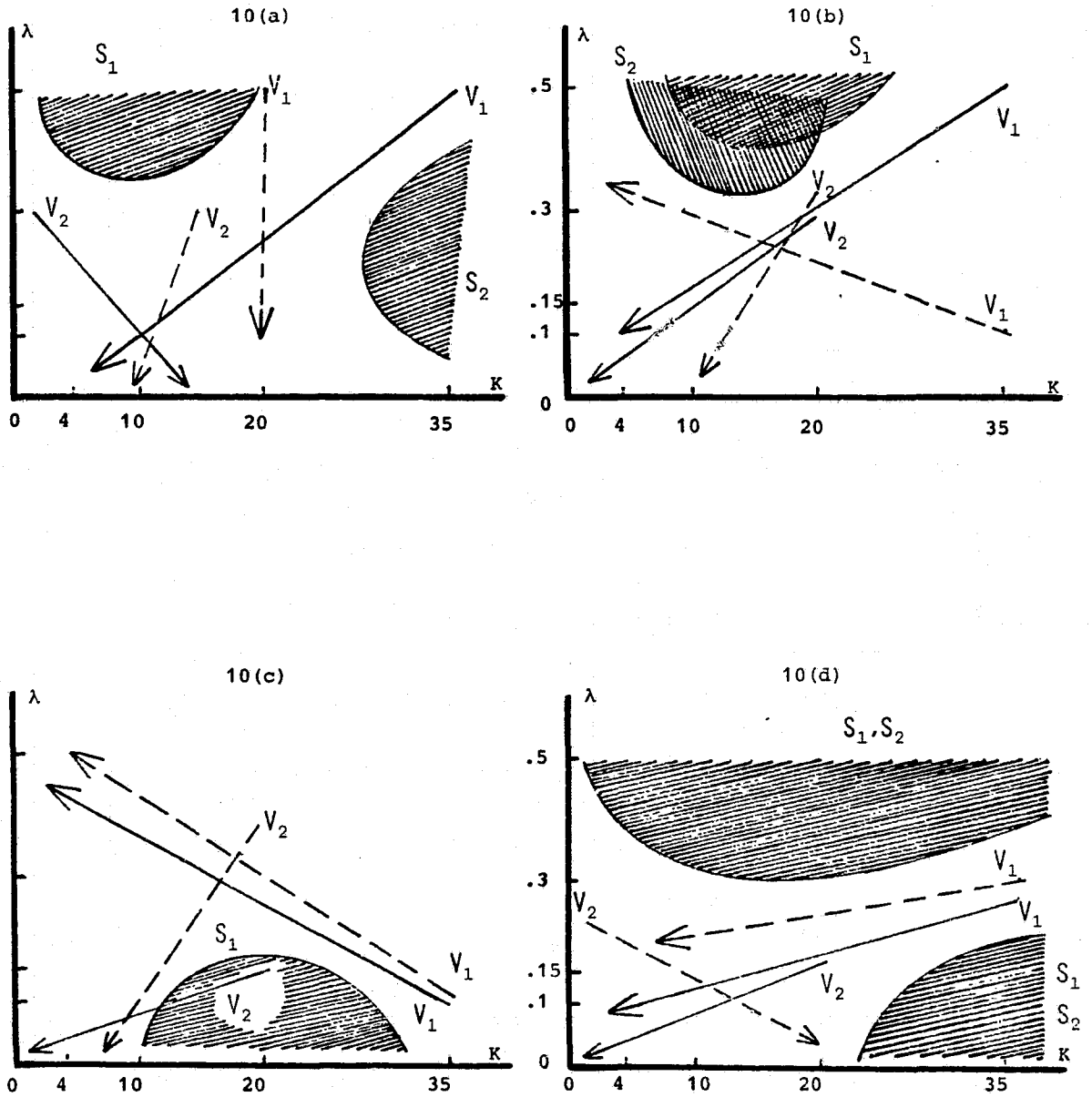
FIGURE 7

SYSTEM PARAMETERS

	1 ..... 5 10
FAMILIARIZATION :	uncomfortable ..... comfortable
DEADAPTATION :	unsatisfactory ..... satisfactory
RELAXATION :	low ..... high
SE / POR :	bad ..... good
FATIGUE :	fair ..... no
EFFORT :	fair ..... no
SIMILARITY :	low ..... high
COMFORT :	low ..... high

FIGURE 8

VERBAL CHARACTERIZATION
(in order of appearance within an experimental session)



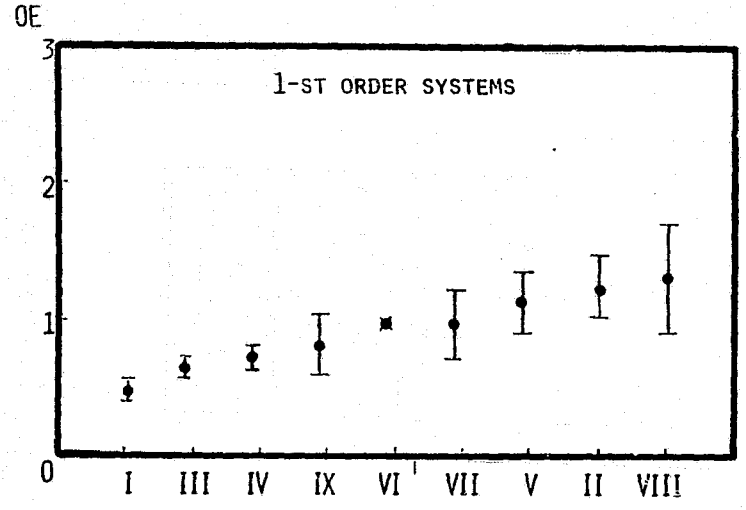
FIGURES 10

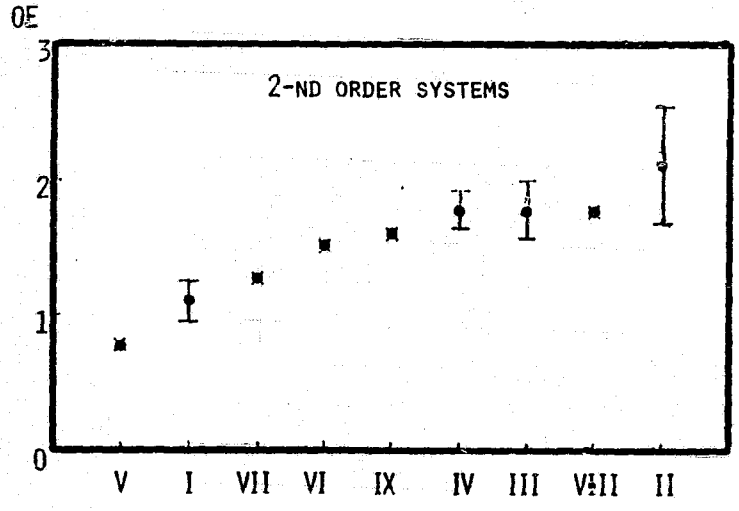
(a) INTRA- AND INTER-RUN RELAXATION VECTORS
(relaxation ——, deadaptation - - - -)

(b) PERFORMANCE AND JUDGEMENT VECTORS
(objective performance ——,
subjective evaluation - - - -)

(c) SUBJECTIVE STRESS STATUS VECTORS
(fatigue ——, effort - - - -)

(d) MEMORIZATION VECTORS
(subjective ——, objective - - - -)



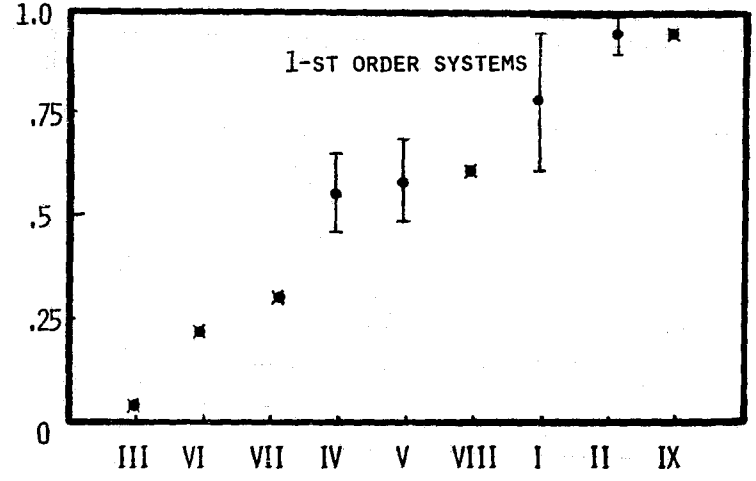


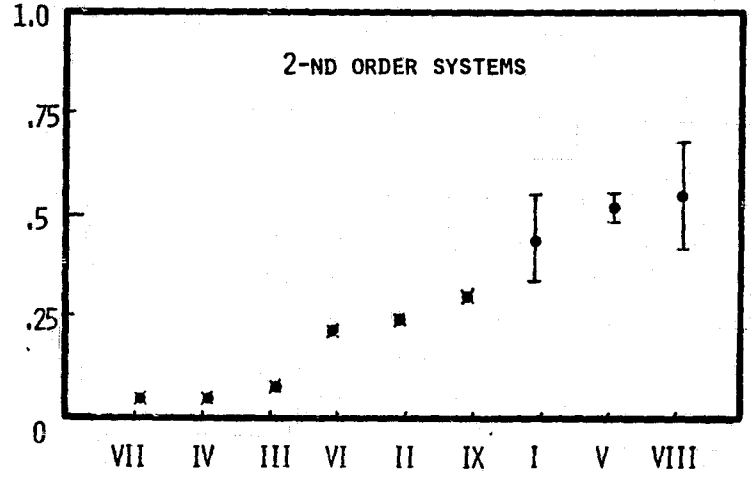
FIGURES 11

PERFORMANCE VARIATIONS

(systems are rank-ordered)

[✱ there is no sufficient sampling to derive valid σ measurements]

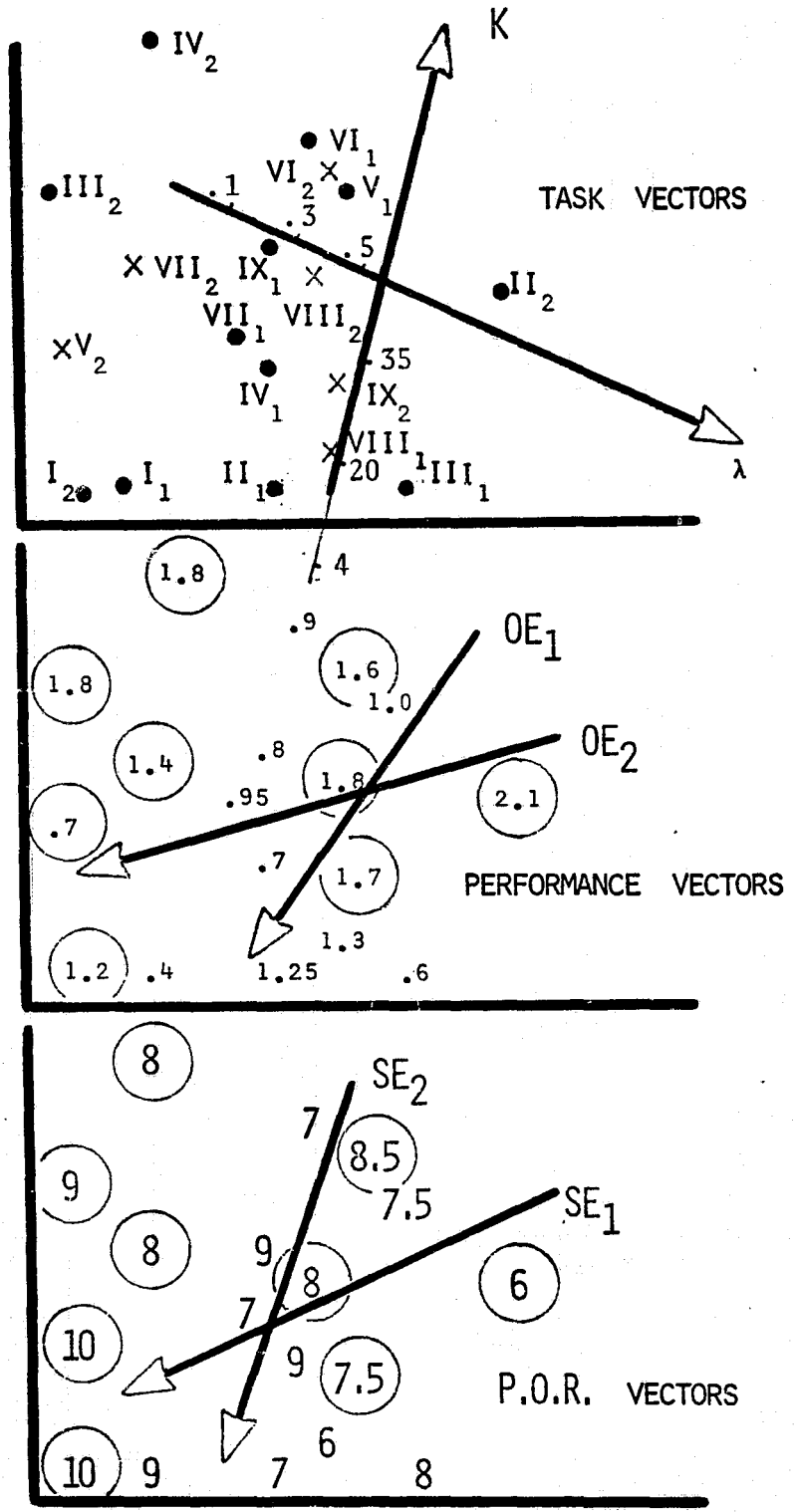




FIGURES 12

SYSTEM IDENTIFIABILITY

(by the H.O.)

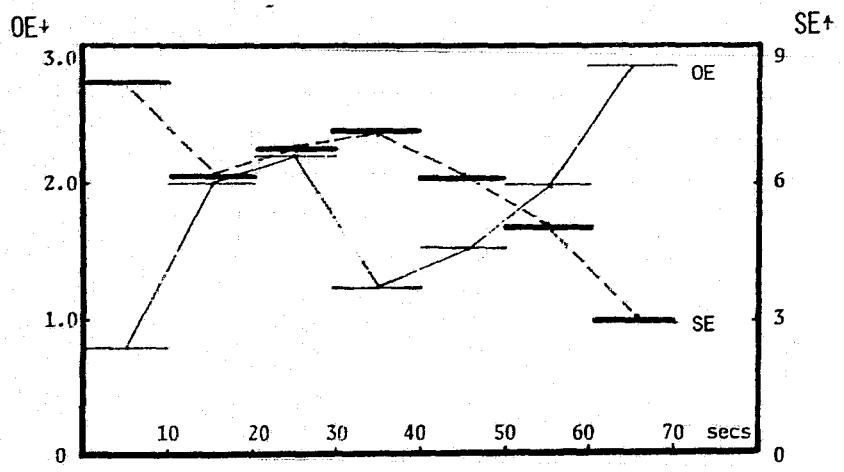


FIGURES 13

PROPERTY VECTORS

(IN THE DISSIMILARITY ITS)

[subscripts signify order of the system; x signify predicted locations of the remaining systems; circled numbers signify 2-nd order systems]



FIGURES 14

VARIABILITY OF TASK II_2

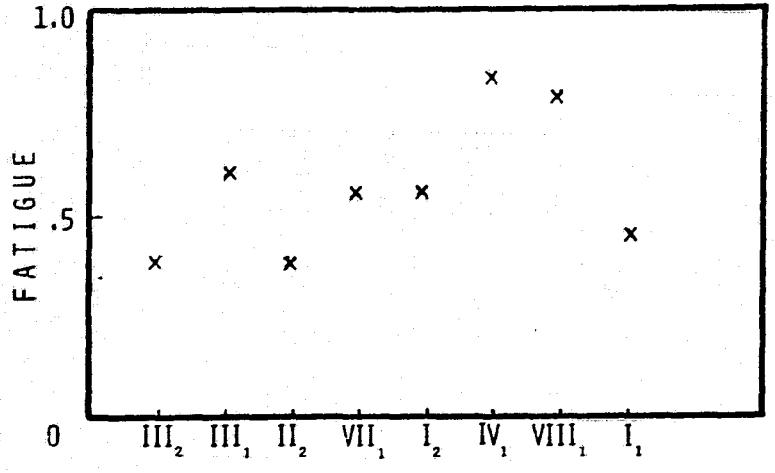
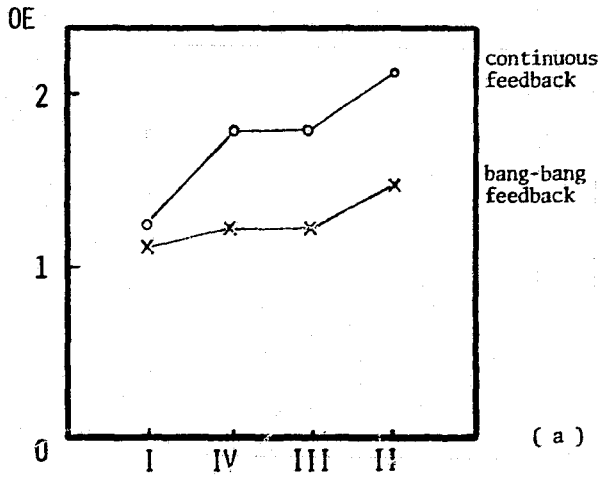
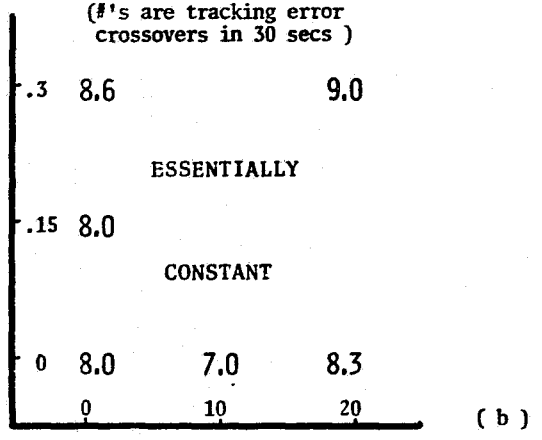


FIGURE 15

IDENTIFIABILITY VS. FATIGUE

(tasks are rank-ordered according to identifiability)





FIGURES 16

BANG-BANG FEEDBACK INFO

(2-nd order systems)

COMPUTATIONAL PROBLEMS IN AUTOREGRESSIVE MOVING AVERAGE (ARMA) MODELS

Gyan C. Agarwal, Shirin M. Goodarzi, William D. O'Neill and Gerald L. Gottlieb

College of Engineering
University of Illinois at Chicago Circle
Chicago, Ill. 60680
and
Department of Physiology
Rush Medical Center
Chicago, Ill. 60612

Abstract

In developing mathematical models of systems from a given input-output data sequence, the choice of the sampling interval and the selection of the order of the model in time-series analysis pose difficult problems. Band-limited (up to 15 Hz) random torque perturbations were applied to the human ankle joint. The applied torque input, the angular rotation output, and the electromyographic activity using surface electrodes from the extensor and flexor muscles of the ankle joint were recorded. Autoregressive moving average models were developed. A parameter constraining technique is applied to develop more reliable models. It is shown that the asymptotic behavior of the system must be taken into account during parameter optimization to develop predictive models.

INTRODUCTION

In a series of previous papers (Agarwal and Gottlieb, 1977 a,b; Gottlieb and Agarwal, 1978; Gottlieb, Agarwal, and Penn, 1978) we have attemped to describe quantitatively the neuromuscluar system dynamics to applied sinusoidal and band-limited gaussian torque perturbations. In these studies, the compliance of the joint was calculated using Fourier series analysis for sinusoidal and power spectral density methods for random perturbations. Although linear analysis methods were used, the system is known to be nonlinear and the parameter values such as the joint viscous and stiffness coefficients are functions of the level of neuromusclar activity.

The purpose of the present paper is to apply time series analysis methods

to study the input-output behavior of the neuromuscular system. The time series method is very parsimonious in the use of parameters to represent the model structure. Normalized residual criterion (NRC) will be used to estimate the model order (For details of this method see Suen and Liu, 1977; Osafo-Charles et. al., 1980).

Our previous analysis was limited to analysis of the angular rotation data and calculation of joint compliance. The electromyographic (EMG) data was not analysed due to inherent difficulties in representing this output by linear transfer functions. The time series approach allows nonlinear representations as long as the model is linear in parameter space.

Dufresne, Soechting and Terzuolo (1978) used pseudo-random torque pulses to study the human forearm response. They developed a model of the EMG in terms of the limb position and its derivatives in the following form:

$$EMG(t) = A\,\theta(t - d) + B\,\dot{\theta}(t - d) + C\,\ddot{\theta}(t - d) \qquad (1)$$

where A, B, and C are constant parameters and d is the time delay. They found that the motor output depends primarily on the angular velocity of the joint. The time delay was found to be about 47 msec.

In a subsquent study, Dufresne, Soechting, and Terzuolo (1979) used different time delay parameters for position and its two derivatives. The best estimates for the time delays were found to be 86 msec for position, 25 msec for velocity, and 45 msec for acceleration. The physiological processes associated with these varying delays are not clear. Soechting and Dufresne (1980) found that the linear model given in equation (1) predicted 80% of the EMG response.

Our analysis of the EMG using time series shows that the autoregressive terms of the EMG are important and cannot be ignored as was done in the Dufresne et. al (1978) model.

METHODS

These experiments were done using normal human subjects. A subject sat in a chair with the right foot strapped to a footplate which could rotate about a horizontal, dorsal-plantar axis through the medial malleolus. The plate could be rotated by a DC torque motor. A band-limited gaussian (0-15 Hz) signal was prerecorded from a noise generator. These time-varying signals were superimposed on a biasing mean motor torque level. The subject was instructed to try to maintain a constant mean force against the bias torque of the motor so that the ankle joint movement was nearly symmetrical with respect to the reference angle. The input was applied for 30 sec or more and the data continuously recorded on a digital tape.

The torque was measured by a strain gauge bridge on the side arms of the footplate. Angular rotation was measured by a continuous capacitive transducer. The EMGs were recorded from disc surface electrodes taped over the bellies of the soleus (SM) and the anterior tibial (TA) muscles. These were amplified full-wave rectified and passed through an averaging filter (10 msec averaging time) before recording. A computer generated the motor drive voltage at a conversion rate of 250/s and digitized data on four input channels. The angle and the torque signals were sampled at a rate of 250/s and filtered EMGs at a rate of 500/s. The data analysis was done off-line using the Minitab 2 statistical software package on an IBM 370 computer.

The Normalized Residual Criterion

Time-series analysis can be extended to obtain discrete linear transfer functions of systems having an input x(t) and output y(t). By x(t) and y(t) we mean pairs of observations that are available at equispaced intervals of time. The behavior of the dynamic system can be adequately represented by the present and past responses and the current and past inputs of the systems. We denote this process as transfer function (TF) models (n,m) and write its equation as

$$y(t) = \alpha_0 + \alpha_1 y(t-1) + \cdots + \alpha_n y(t-n) + \beta_0 x(t) + \ldots + \beta_n x(t-n) + \nu(t) \quad (2)$$

In (2) the parameters to be estimated are $\alpha_0,\ldots,\alpha_n$, $\beta_0,\ldots,\beta_m$, n, and m. The time series $\nu(t)$ is a random term measuring the difference between the response y(t) and the variables used to expalin the time-series data. The parameter α_0 measures the mean output.

Equation (2) reduces to an autoregressive model (AR(n)) if x(t) is omitted from the model, and reduces to a moving average model (MA(m)) if lags of y are omitted. The following assumptions will be made concerning $\nu(t)$ for a given ouput time sequence y(t), t = [0,T],

1) $E[\nu(t)] = 0$

2) $E[\nu(i)\nu(j)] = \sigma_\nu^2 \delta_{ij}$ (3)

where

$$\delta_{ij} = \begin{cases} 1 \text{ for } i=j \\ 0 \text{ for } i \neq j \end{cases}$$

3) $T >> n$.

From (2),

$$\nu(t) = y(t) - \alpha_0 - \sum_{i=1}^{n} \alpha_i y(t-i) - \sum_{j=0}^{m} \beta_j x(t-j),$$

$$t = 1, 2, \cdots, T-n. \quad (4)$$

Define

$$\sum_{t=1}^{T-n} \nu^2(t) = ||V||^2 \text{ and } \sum_{t=1}^{T-n} y^2(t) = ||Y||^2. \quad (5)$$

Note that in the discussion below V and Y are vectors such that

$$V=\begin{bmatrix} \nu(1) \\ \nu(2) \\ \vdots \\ \nu(T-n) \end{bmatrix} \qquad Y=\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(T-n) \end{bmatrix} \tag{6}$$

Squaring (4) and normalizing by the total sum of squares, we have

$$\frac{||V||^2}{||Y||^2} = \frac{||y-\alpha_0-\Sigma\alpha_i y - \Sigma\beta_j x||^2}{||Y||^2} = \varepsilon(n, m, T) \tag{7}$$

and therefore

$$E\left[\, ||V||^2 \,\right] = E\left[\, ||Y||^2 \varepsilon(n,m,T) \right] . \tag{8}$$

Since y(t), the data series, is deterministic, (8) can be rewritten as

$$E\left[\, ||V||^2 \,\right] = ||Y||^2 E\left[\, \varepsilon(n, m, T) \,\right]. \tag{9}$$

From (6) we have

$$E\left[\, ||V||^2 \,\right] = E\left[\nu^2(1) + \nu^2(2) + \cdots + \nu^2(T-n)\right] , \tag{10}$$

which by assumption 2) in (3) reduces to

$$E\left[\, ||V||^2 \,\right] = \left[T-n\right]\sigma_\nu^2 . \tag{11}$$

Substitution of (11) in (9) we have

$$E\left[\varepsilon(n,m,T)\right] = \frac{(T-n)\sigma_\nu^2}{||Y||^2} \tag{12}$$

and by assumption 3), (12) becomes

$$E\left[\varepsilon(n, m, T)\right] \simeq \frac{T\sigma_\nu^2}{||Y||^2} \tag{13}$$

The quantity $\varepsilon(n,m,T)$ depends on n, m, and T and is proportional to the normalized variance of the regression for a given n and m. If this ratio is minimized over n and m, then the data fit as measured by the correlation coefficient ρ will be maximized. Note that

$$\rho = \left[1 - \frac{||V||^2}{||Y||^2} \right]^{1/2} \tag{14}$$

or

$$\hat{\rho} = \left[1 - \hat{\varepsilon}(n,m) \right]^{1/2} \tag{15}$$

where T, being a constant for the data, is omitted in the optimization procedure, and $\hat{\varepsilon}(n,m)$ is the minimum value for $\varepsilon(n,m)$. This optimization technique is the so called the Normalized Residual Criterion.

RESULTS

The mathematical modeling problem was considered in two separate parts. For the first model the applied torque is the input and the resulting angular rotation of the joint is the output of the system. For the second model, the angular rotation (and its derivatives) is considered as the input to the system and the resulting stretch reflex electromyographic activity is considered as the output. It should be emphasized that the angular rotation is the net result of two torque inputs applied at the joint; one by the external motor torque and the other muscle forces produced by the stretch reflex mechanism. These mechanisms are also responsible for a significant contribution to the joint viscous and elastic properties. Figure 1 shows a sample of the data at 4 msec sampling interval. The velocity was obtained by digital differentiation.

Angle-Torque Model

Although the data was recorded for 30 seconds at each input (Agarwal and Gottlieb, 1977b), this method does not require such long data records which would also use too much computer time. The time series analysis was done using only two-seconds of the data record. (The first two-seconds of the data were not used to allow the turn-on transients to die out).

The values of $\varepsilon(n,m)$ were computed for a given data record and then plotted against different values of n (see Figure 2). The data sampling interval in this case is 4 msec. This analysis clearly indicates that $n = 2$ and $m = 0$ is adequate to model this data. The same data was analyzed again using the sampling intervals of 12, 20, 40, and 60 msec. Figure 3 shows the $\varepsilon(n,m)$ values for the sampling interval of 20 msec. Note that the minimum value of the normalized residual is about 60 times of that in the first case. For 40 and 60 msec sampling $\varepsilon(n,m)$ did not reach asymptotic values even for model order of (8,8). The normalized residuals at 12 and 20 msec sampling implied a model order of (3,1).

Figure 4 shows the actual angular rotation data (2 to 4 sec interval used in this analysis), the regression fit and the predicted output using 4 msec sampling and model order of (3,1). The regression fit is obtained by using the equation

$$\theta(t) = \alpha_0 + \alpha_1\,\theta(t-1) + \alpha_2\,\theta(t-2) + \alpha_3\,\theta(t-3) + \beta_0 T(t) + \beta_1 T(t-1) \tag{16}$$

The error between the actual data and the regression fit is nearly zero. The correlation coefficient is $\rho = 0.999$. However, when this model is used to predict the output using the first three data output values as the initial conditions, the predicted output is a poor approximation of the actual data (see Figure 4). Figure 5 shows the observed angle and the predicted model values for model orders of (3,1), (7,1), (9,1), and (14,1). Even the fourteenth order model is not able to adequately reproduce the data sequence. These models are not able to capture the steady state (or long term) behavior of the system. Osafo-Charles, et al., (1980) showed that to develop better predictive models, the TF(n,m) models must be constrained to incorporate the steady state response of the system.

Constrained Model

Consider the estimated model given by equation (16). Under conditions of equilibrium

$$\theta(t) = \theta(t-1) = \theta(t-2) = \theta(t-3) = \theta_e$$

and

$$T(t) = T(t-1) = T_e$$

where θ_e and T_e are the steady state response and input respectively. At physical and statistical equilibrium, with $\theta(t) = \theta_e$ and $T(t) = T_e$, equaition (16) became

$$\theta_e = \alpha_1\,\theta_e + \alpha_2\,\theta_e + \alpha_3\,\theta_e + \beta_0\,T_e + \beta_1\,T_e \tag{17}$$

or

$$\frac{\theta_e}{T_e} = \frac{\beta_0 + \beta_1}{1 - \alpha_1 - \alpha_2 - \alpha_3} = g \tag{18}$$

where (18) expresses the steady state gain in terms of the parameters of the model.

The value of g was approximated by the slope of the curve of torque vs. angular rotation in the relaxed ankle during sinusoidal oscillation at 0.1 H_z (Gottlieb and Agarwal, 1978).

For $\theta_e = g\ T_e$ to be true, we must have

$$\beta_0 = g(1 - \alpha_1 - \alpha_2 - \alpha_3) - \beta_1 \tag{19}$$

From equations (19) and (16), we get

$$\begin{aligned}\left[\theta(t) - g\ T(t)\right] = {} & \alpha_1 \left[\theta(t-1) - g\ T(t)\right] \\ & + \alpha_2 \left[\theta(t-2) - g\ T(t)\right] \\ & + \alpha_3 \left[\theta(t-3) - g\ T(t)\right] \\ & + \beta_1 \left[T(t-1) - T(t)\right]\end{aligned} \tag{20}$$

Regression analysis is used again to estimate the parameters α_1, α_2, α_3 and β_1 for a given value of gain g. β_0 is then obtained using (19). Figure 6 shows the output angle and predicted model response for a constrained model with gains of g = 6.5, 7.5, and 8.5. The gain value of 8.5 was considered to provide the best fit in terms of the minimum estimated standard deviation of the regression.

The transfer function for the unconstrained model is:

$$H(z) = \frac{0.00239 - 0.00024\ z^{-1}}{1 - 2.678\ z^{-1} + 2.399\ z^{-2} - 0.7191\ z^{-3}} \tag{21}$$

For the constrained model with a slope of 8.5, the transfer function is:

$$H(z) = \frac{0.00238 + 0.00017\ z^{-1}}{1 - 2.731\ z^{-1} + 2.503\ z^{-2} - 0.7717\ z^{-3}} \tag{22}$$

EMG Model

Our efforts to model EMG as a function of either the angular rotation or the velocity or a combination of both were not successful. As was noted by Dufresne et al. (1978), the velocity of rotation is the most significant input due to spindle properties (Matthews, 1972). However only those components of velocity which stretch the spindle contribute to the EMG of the stretched muscle. (The spindle is silent during shortening). Therefore, a new velocity signal represneting only the stretching velocities was defined as:

$$\dot{\theta}_d(t) = \dot{\theta}(t) \quad \text{if} \quad \dot{\theta} \geq 0$$
$$= 0 \quad \text{if} \quad \dot{\theta} < 0 \tag{23}$$

The normalized residual analysis indicated a model order of (4,1) using soleus EMG as the ouput and $\dot{\theta}_d$ as the input signal. The predicted output of the unconstrained model and its comparison with the actual EMG signal is shown in Figure 7. Since the EMG signal is a full-wave rectified and filtered (using an averaging filter) signal, it has only negative values (because of negative filter gain). The predicted value of EMG is a poor approximation of the data.

A constrained model was developed using a similar approach as outlined earlier. Figure (8) shows the predicted EMG and the actual data at three values of the gain parameter. The gain of -0.005 was considered to give the most appropriate fit.

For the constrained model with a slope of -0.005, the transfer function is:

$$H(z) = \frac{EMG}{\dot{\theta}_d} = \frac{-.004164 + .00314\ z^{-1}}{1 - 1.19\ z^{-1} + .6685\ z^{-2} - .2947\ z^{-3} + .02104\ z^{-4}} \tag{24}$$

CONCLUSIONS

The time series approach is a powerful and versatile technique in developing time domain models from a given input-output data sequence. Normalized residual criterion allows effective prediction of the model order. Models developed in this manner may be satisfactory, but may not be good predictive models. It is recommended that constrained parameter modeling which allows incorporating the steady-state behavior be used to obtain better predictive models.

Acknowledgment

This work was supported in part by National Science Foundation grant ENG-7608754 and grants from the National Institutes of Health NS-12877 and NS-00196.

REFERENCES

1. G. C. Agarwal and G. L. Gottlieb. Oscillation of the human ankle joint in response to appled sinsoidal torque on the foot. J. Physiology (London), 268:151-176, 1977a.

2. G. C. Agarwal and G. L. Gottlieb. Compliance of the human ankle joint. ASME J. of Biomechanical Engineering, 99:166-170, 1977b.

3. J. R. Dufrense, J. F. Soechting, and C. A. Terzuolo. Electromyographic response to pseudo-random torque disturbances of human forearm position. Neuroscience, 3:1213-1216, 1978.

4. J. R. Dufrense, J. F. Soechting, and C. A. Terzuolo. Reflex motor output to torque pulses in man: Identification of short - and long- latency loops with individual feedback parameters. Neuroscience, 4:1493-1500, 1979.

5. G. L. Gottlieb, G. C. Agarwal, and R. Penn. Sinusoidal oscillation of the ankle as a means of evaluating the spastic patient. J. Neurology, Neurosurgery, and Psychiatry, 41: 32-39, 1978.

6. G. L. Gottlieb and G. C. Agarwal. Dependence of human ankle compliance on joint angle. J. Biomechanics, 11:177-181, 1978.

7. P. B. C. Matthews. Mammalian Muscle Receptors and Their Central Actions, Williams & Wilkins Comp., Baltimore, 1972.

8. F. Osafo-Charles, G. C. Agarwal, W. D. O'Neill, and G. L. Gottlieb. Application of time-series modeling to human operator dynamics. IEEE Trans. Systems, Man, and Cybernetics, SMC-10:849-860, 1980.

9. J. F. Soechting and J. R. Dufresne. An evaluation of nonlinearities in the motor output response to applied torque perturbations in man. Biol. Cybernetics, 36:63-71, 1980.

10. L. C. Suen and R. Liu. A normalized residual criterion for the determination of structure of linear systems. Policy Analysis & Information Systems, 1:1-22, 1977.

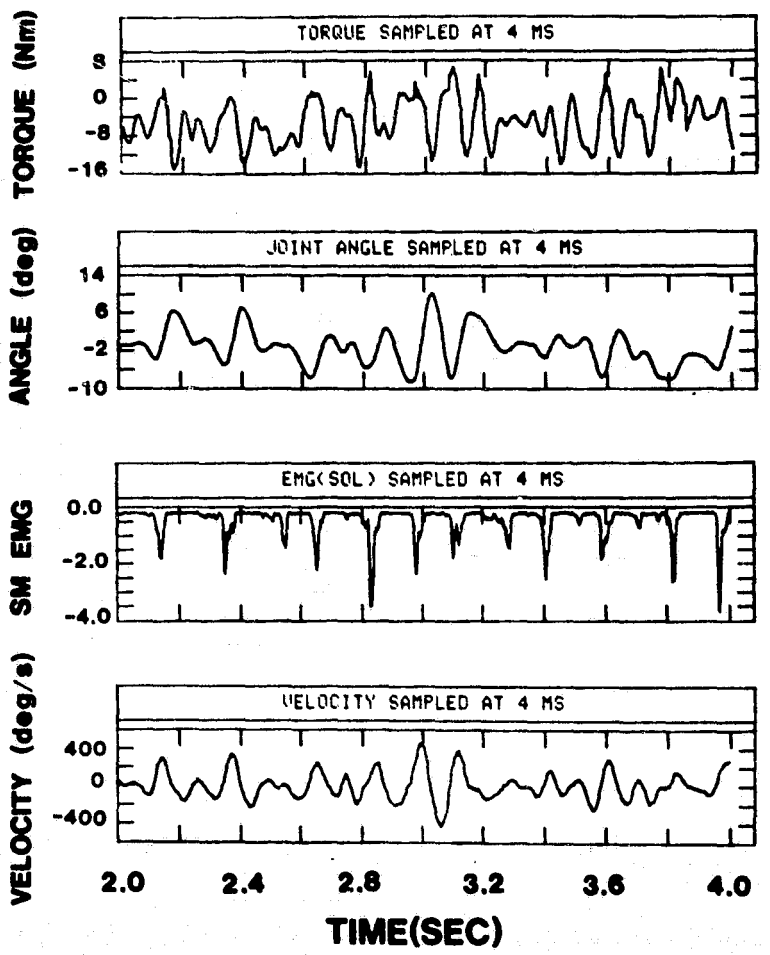

TORQUE SAMPLED AT 4 MS
TORQUE (Nm)
8
0
-8
-16
JOINT ANGLE SAMPLED AT 4 MS
ANGLE (deg)
14
6
-2
-10
EMG(SOL) SAMPLED AT 4 MS
SM EMG
0.0
-2.0
-4.0
VELOCITY SAMPLED AT 4 MS
VELOCITY (deg/s)
400
0
-400
2.0
2.4
2.8
3.2
3.6
4.0
TIME(SEC)

Figure 1

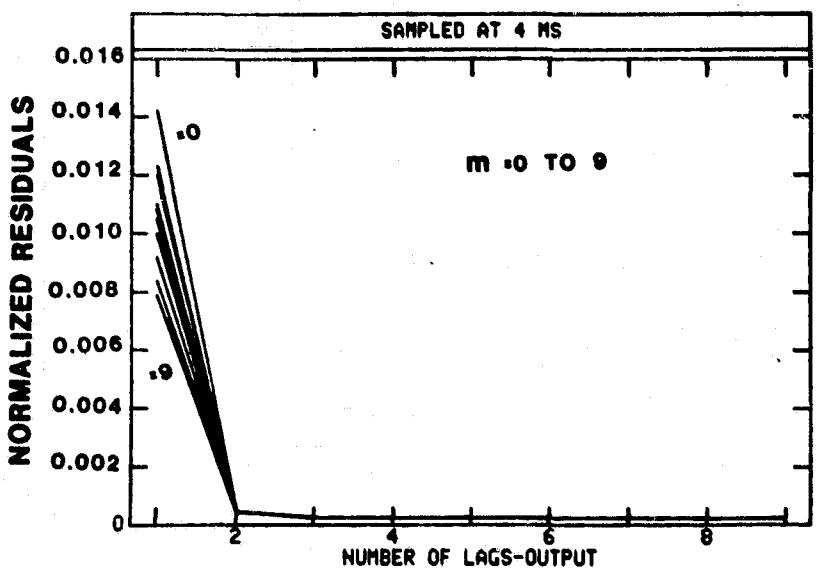


Figure 2

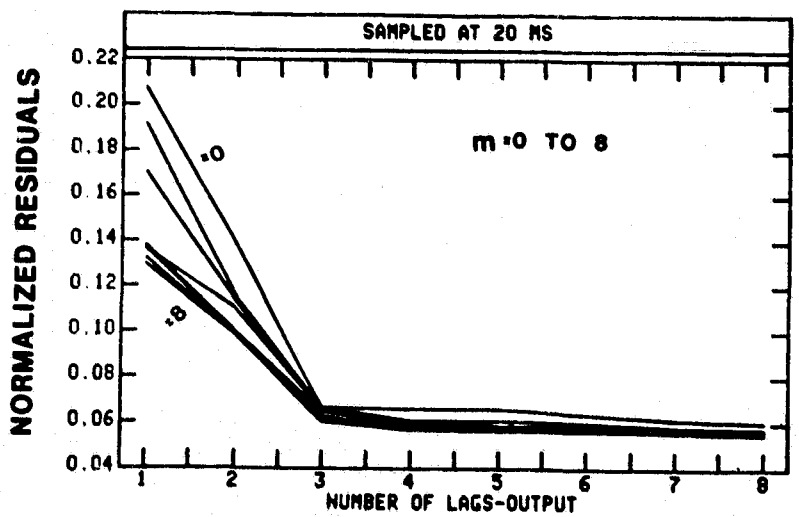


Figure 3

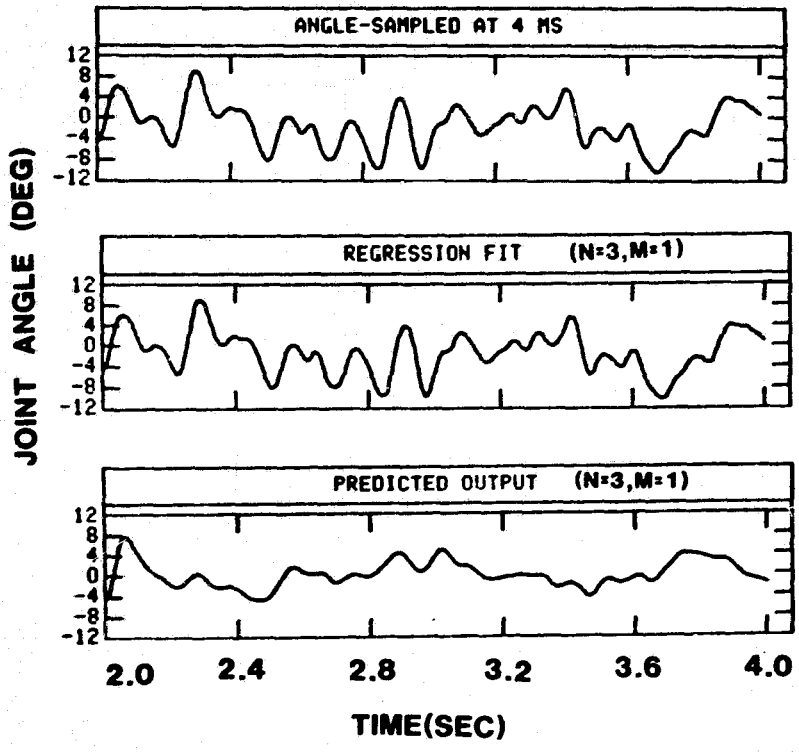


Figure 4

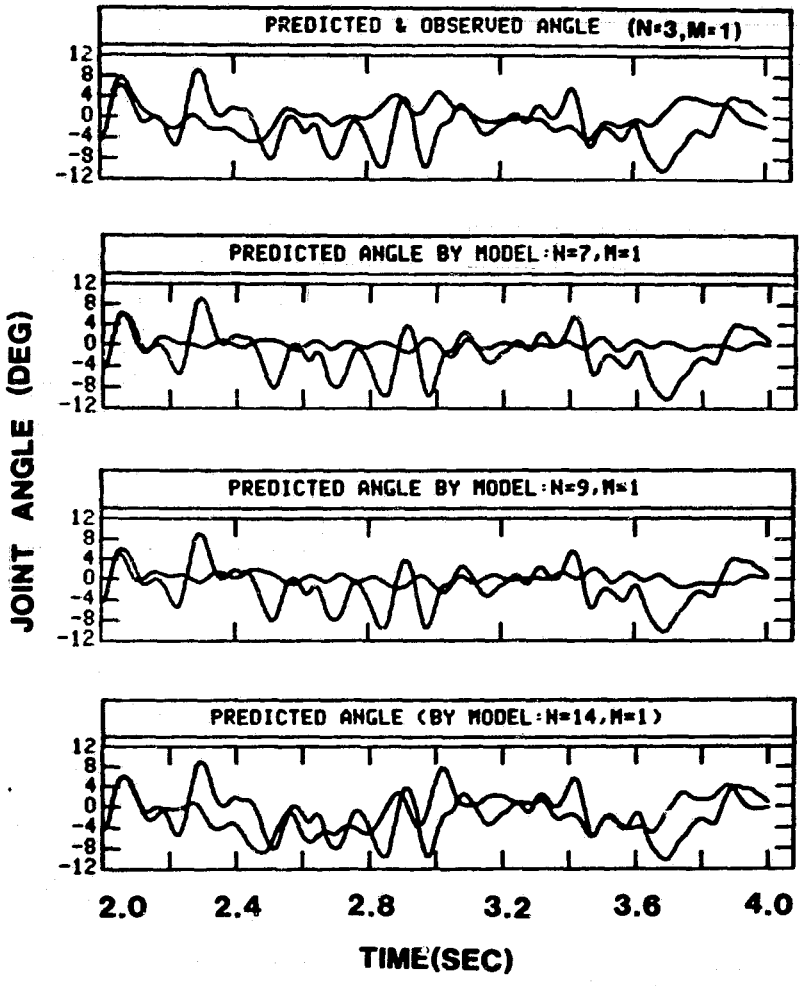


Figure 5

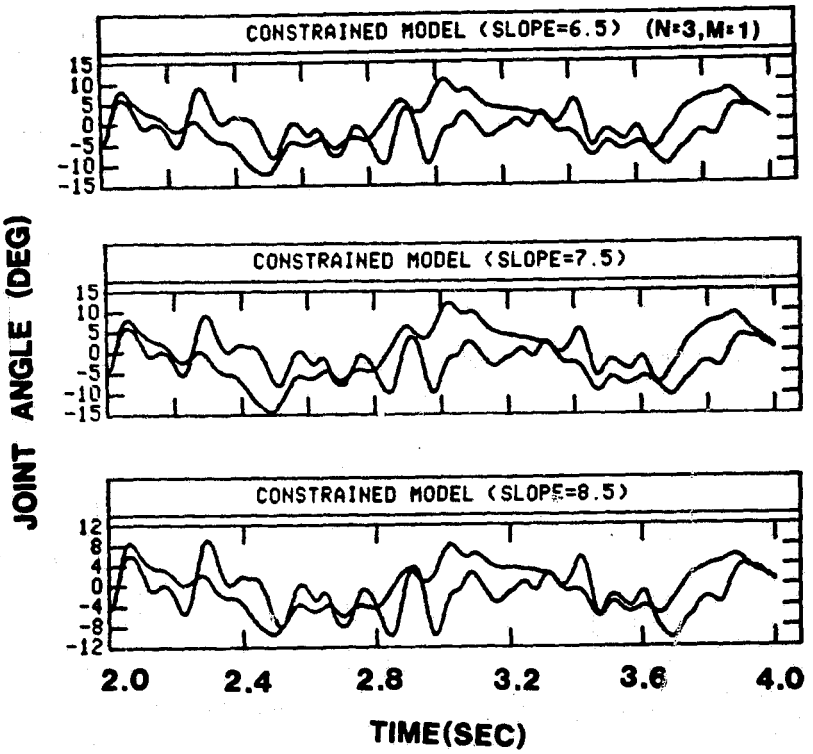


Figure 6

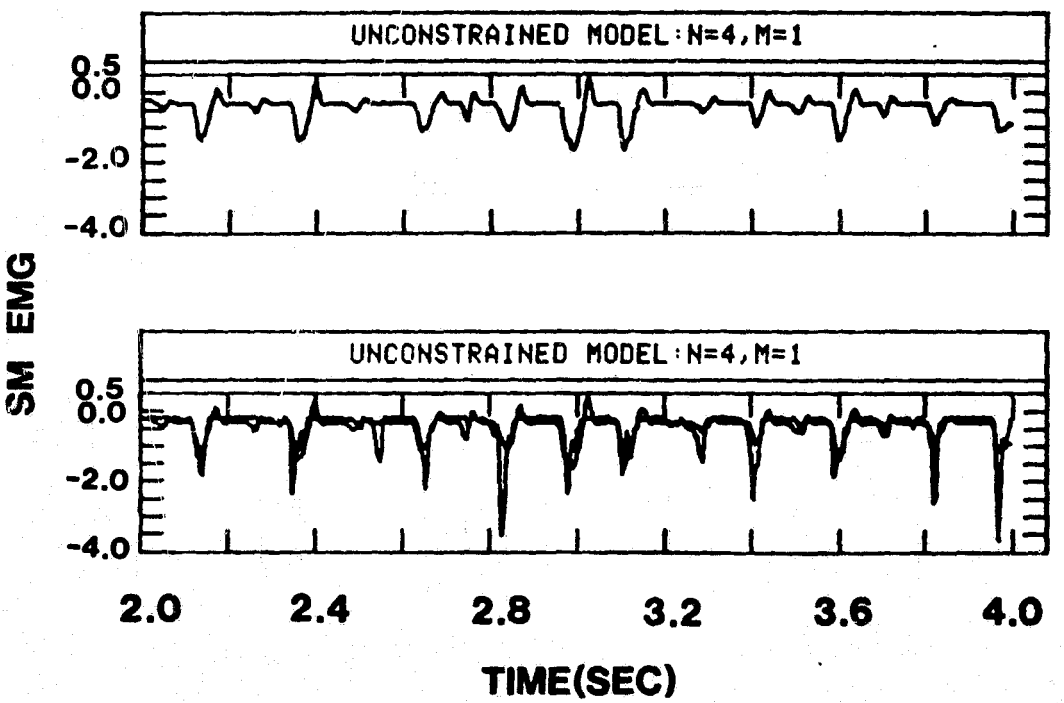


Figure 7

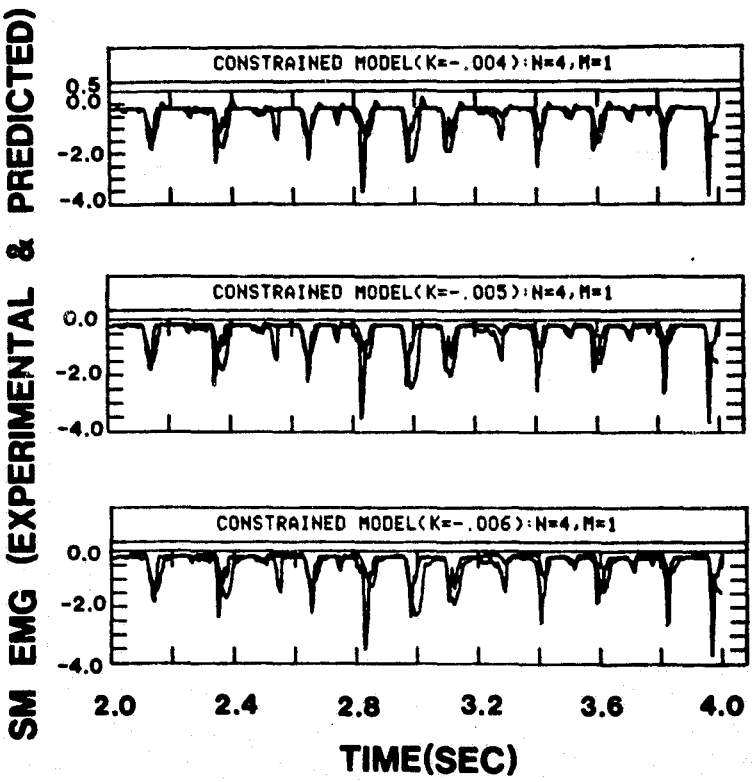

CONSTRAINED MODEL(K=-.004):N=4,M=1
CONSTRAINED MODEL(K=-.005):N=4,M=1
CONSTRAINED MODEL(K=-.006):N=4,M=1
0.5
0.0
-2.0
-4.0
2.0
2.4
2.8
3.2
3.6
4.0
SM EMG (EXPERIMENTAL & PREDICTED)
TIME(SEC)

Figure 8

D34

N82-13699

MODELING HUMAN TRACKING ERROR IN SEVERAL DIFFERENT ANTI-TANK SYSTEMS†

David L. Kleinman

ALPHATECH, Inc. and University of Connecticut

SUMMARY

The Optimal Control Model (OCM) of human response serves as a mechanism for generating sample time histories of human tracking error in different anti-tank systems. The systems under study include TOW (Tube-Launched Optically Guided System), DRAGON (Shoulder Mounted) and ITV (Improved TOW Vehicle). The model-generated trajectories are compared with field-test data across several dimensions including time-domain (temporal) statistics, frequency content and subjective comparisons on individual runs.

MODELING APPROACH

The objective of this work is to develop a computerized model for generating human tracking error time histories in several different anti-tank systems. The systems include those in common use by the US Army such as TOW, DRAGON and ITV. A fourth system - GLLD (Ground Launched Laser Designator) is similar to TOW and will not be discussed here. Of these systems, TOW and DRAGON are basically command line-of-sight (LOS), whereas ITV is a rate command system. The model that is developed must produce accurate facsimiles of tracking error over a wide range of target trajectories, from crossing (straight-line) motion to maneuvering motion. The model must be causal in the sense that future target motions are unknown at the present time.

For the systems and target passes considered here, target motion is restricted to the azimuth axis, i.e. the gunner and target vehicle are both at the same ground level. Tracking error in elevation arises solely from the human's inherent motor and observation randomness. Although the model that we have developed treats both axes, we discuss primarily the results for the azimuth axis here. A more complete discussion and presentation of the results may be found in Ref. [1].

The Optimal Control Model

The Optimal Control Model of human response is used as the mechanism for building the anti-tank tracking model. The OCM technology has been successfully applied in numerous contexts including pilot control, anti-aircraft artillery, etc. The mechanics of using the OCM to generate sample path time-histories (as opposed to statistical measures) is described in Ref. [2]. Our application follows this approach, with minor modifications to account for the dual-axis nature of the tracking task. The pertinent

†Supported by Army Material Systems Analysis Agency on Contract DAAK11-80-C-0050

equations for Monte-Carlo/Sample Path Simulation using the OCM may be found in Ref. [2].

Application of the OCM requires specification, for each given system-display-manipulator dynamics, of 1) the operator's task objectives in terms of a set of cost functional weights, 2) the parameters that define the operator's inherent limitations, and 3) an "internal" model of the target dynamics. With these items specified optimal control and estimation theory is used to obtain the human's feedback strategy, and generate closed-loop performance results.

1. *Task Objectives:* For a basic tracking task, wherein the human attempts to keep the error e(t) small, we use a cost functional

$$J(u) = E\{e^2(t) + Q_{\dot{e}}\, \dot{e}^2(t) + Q_{\dot{u}}\, \dot{u}^2(t)\} \qquad (1)$$

The weighting $Q_{\dot{e}}$ reflects a human's subjective weighting on error rate. It is indicative of strategy, style or technique and could be associated with the type of training on a given system. The weighting $Q_{\dot{u}}$ on error rate induces a first-order lag that is associated with the neuro-motor system dynamics. For each system we select $Q_{\dot{u}}$ to yield τ_n=neuro-motor time-constant = .1 sec. The value of $Q_{\dot{e}}$ is to be determined for each system, based on data comparisons.

2. *Human Operator Limitations:* The primary human operator limitations modeled in the OCM are those associated with perceiving displayed quantities and executing intended control motions. The observational submodel in the OCM assumes that the human observes tracking error $y_1(t) = e(t)$ and tracking error rate $y_2(t) = \dot{e}(t)$. However, the human perceives a delayed and noisy replica of these signals via

$$y_{pi}(t) = F_i[y_i(t - \tau)] + v_{yi}(t - \tau) \qquad i = 1,2 \qquad (2)$$

The function $F(\cdot)$ represents a visual/indifference threshold of value a_1 = 0.6 mr ÷ display gain on error and $a_2 = .5a_1$ on error rate. The time-delay τ = .15 sec. Each observation noise is white with covariance

$$V_{yi}(t) = \rho_{yi}\, E\{y_i^2(t)\}/\ ATTN \qquad (3)$$

Where $\rho_{yi} = .01\pi$ (-20dB); the attention allocations are assumed split 0.8 for azimuth vs 0.2 for elevation.† The nominal parameter values associated with the observational submodel are assumed fixed for all systems and target types considered.

The neuro-motor submodel for generating human corrective inputs is given by

$$\tau_n\, \dot{u}(t) + u(t) = u_c(t) + v_u(t)$$

†A more precise model would be to employ dynamic attention allocation.

The quantity $u_c(t)$ is the "commanded" control input that is generated from the Kalman Filter/Predictor/Gains cascade; τ_n is the neuro-motor time constant. The white motor-noise $v_u(t)$ consists of an additive plus a ratioed component for each axis. The covariance of the motor-noise for azimuth and elevation axes, respectively, is

$$V_{uA} = V^o_{uA} + \rho_{AA}\ E\{u^2_{cA}(t)\} + \rho_{AE}\ E\{u^2_{cE}(t)\} \tag{4a}$$

$$V_{uE} = V^o_{uE} + \rho_{EA}\ E\{u^2_{cA}(t)\} + \rho_{EE}\ E\{u^2_{cE}(t)\} \tag{4b}$$

The crossfeeds ρ_{AE} and ρ_{EA} model uncertainty/randomness in one axis resulting from manipulator motion in the other axis. For the systems/targets studied, elevation commands u_{cE} are small relative to azimuth commands u_{cA} (recall target motion is in azimuth only). Thus, ρ_{EE} and ρ_{AE} are not readily obtained from the available data. Therefore, we have assumed

$$\rho_{AA} = \rho_{EE} = \rho_u \quad \text{and} \quad \rho_{AE} = 0 \tag{5}$$

The remaining quantities V^o_{uA}, V^o_{uE}, ρ_u, ρ_{EA} are system/manipulator dependent. Their values must be elicited from model - data comparisons. Finally, Eq(4) shows that the motor-noise scales with commanded control input $u_c(t)$. In some instances, e.g. command LOS systems, it is more natural for motor-induced randomness to scale with commanded rate $\dot{u}_c(t)$.

3. *Target Submodel:* In the present application of the OCM to track ground targets, target velocity $\dot{\theta}_T(t)$ and acceleration $\ddot{\theta}_T(t)$ are generally small. Thus, we use a simple internal model for target motion

$$\dot{x}(t) = w_d(t) \tag{6}$$

where $x(t)$ is the human's internal representation of target velocity. The "driving noise" $w_d(t)$ has covariance

$$\text{cov}\ [w_d(t)] = \beta\ \ddot{\theta}^2_T(t) + \alpha\ \dot{\theta}^2_T(t) \tag{7}$$

Note that the "truth" model is $\dot{x}(t) = \ddot{\theta}_T(t)$. The values selected for α and β are

$$\beta = 10^2, \quad \alpha = .015^2 \tag{8}$$

These values are constant across all systems and targets studied.

Data-Model Comparison Procedures

The OCM can be used to generate, for a given system and target trajectory, an ensemble of tracking error time histories $E_m = \{e_j(t)\ ;\ j = 1,...M\}$. These model-generated runs may be compared against an ensemble of equivalent data trials, E_d. Clearly, it is the statistics of these two ensembles that one would wish to compare via model-data validation tests. Several modes of comparison are possible, as discussed below. For consistency we have found it useful to remove the temporal mean μ_{ej} from each run prior to analysis.

Thus,

$$e_j(t_i) = e_j(t_i) - \mu_{ej} \; ; \; \mu_{ej} = \frac{1}{N} \sum_{i=1}^{N} e_j(t_i) \tag{9}$$

where t_i, $i = 1,...N$ are the sampled values of $e_j(t)$. This procedure removes random variations in signal mean, DC bias offsets in field recording equipment, biases in human aiming point, etc.

1. *Ensemble Analysis:* The ensemble mean and standard deviation can be computed in the usual manner,

$$\mu_e(t) = \text{Ensemble mean} = \frac{1}{M} \sum_{j=1}^{M} e_j(t) \tag{10a}$$

$$\sigma_e(t) = \text{Ensemble SD} = \left\{ \frac{1}{M-1} \sum_{j=1}^{M} [e_j(t) - \mu_e(t)]^2 \right\}^{1/2} \tag{10b}$$

Since the sample runs have been rendered zero (temporal) mean we would expect $\mu_e(t) \approx 0$ if the ensemble was stationary.

2. *Temporal Analysis:* If the ensemble E is stationary, $\mu_e(t) \approx 0$ and $\sigma_e(t) \approx$ constant so that the information content in the temporal ensemble is reduced to a single number, i.e. RMS tracking error. A more direct way to obtain average RMS tracking error is to compute the temporal statistic

$$\sigma_{ej} = \left[\frac{1}{N-1} \sum_{i=1}^{N} e_j^2(t_i) \right]^{1/2} \tag{11}$$

for each run, and form the composite, M-run average, via

$$\sigma_e = \frac{1}{M} \sum_{j=1}^{M} \sigma_{ej} \tag{12}$$

However, unlike the ensemble analysis, σ_e is meaningful only in the stationary case. Its computation in the non-stationary case is possible, but of dubious interpretation.

3. *Frequency Domain Analysis:* The RMS temporal metrics give an indication of total error power, they do not indicate how this power is distributed over the frequency range, whether there exists resonances, etc. To obtain these later indicators of system response we compute, from the temporal ensemble $E = \{e_j(t), j = 1,...,M\}$, a frequency domain ensemble of normalized error PSD, $E^* = \{e_j^*(\omega), j = 1,...,M\}$ where

$$e_j^*(\omega) = \left| \frac{1}{\sigma_{ej}} \sum_{i=1}^{N} e_j(t_i) \exp[-j\omega t_i] \right| \tag{13}$$

Normalization greatly reduces the sensitivity to motor-noise and inter-subject variability since it considers only relative power distribution over ω. Note that the PSD computations are strictly valid for a sta-

tionary ensemble; their computation is possible for any ensemble, of course, but the interpretation in the non-stationary case is dubious.

The ensemble E^* of normalized error PSD can be averaged (in the same manner as E) via equations similar to (10a - 10b) to yield the ensemble mean $\mu_e^*(\omega)$ and ensemble SD, $\sigma_e^*(\omega)$. Comparison of model and data PSD statistics is thus possible and provides another, interesting, facet for model validation.

RESULTS

In this section model-data comparisons are given for the three anti-tank systems considered. In each case it is necessary to provide a description of the system-manipulator dynamics, and values for the motor-noise parameters and error rate weighting Q_e.

TOW System

The TOW system is a command LOS system consisting of a launch tube plus sight mounted on a viscous (rate) damped turret. Thus the torque supplied by the operator to point the sight varies with sight (i.e. control) rate. The dynamical model used for the TOW system is

$$T(s) = \frac{1}{Ts+1} \quad ; \quad T = 0.1 \text{ sec.} \tag{14}$$

These dynamics are chosen for convenience†, and are viewed as representative of the manipulator (arm-viscous mount) characteristics.

In our data-model analysis of the TOW system we found $Q_e \simeq 0$ gave best match. The motor-noise in the OCM is assumed to scale with commanded (i.e. LOS) rate, and the pertinent noise parameters (obtained from matching RMS scores) are

$$V_u^o = [.05, .01], \; \rho_u = .005, \; \rho_{EA} = .015$$

The temporal RMS statistics, computed via Eqs (11)-(12), for three different ensembles are given in Table 1. The TF and TS ensembles correspond to crossing targets at a range of 3Km with $\dot{\theta}_T = 5.47$ and $\dot{\theta}_T = .55$ mr/sec, respectively. These ensembles are stationary. The TM ensemble corresponds to a set of 19 maneuvering trials. In these cases the target was approaching the gunner following a serpentine path with $\dot{\theta}_T \sim 1$ mr/sec, $\ddot{\theta}_T \sim .5$ mr/sec^2 peak values. Each target pass was somewhat different; this ensemble is not stationary.

The model-data comparisons shown in Table 1 are excellent for both azimuth and elevation axis tracking. (The numbers in parentheses are the computed standard deviations in the RMS tracking errors.) Only the elevation SD is not well-matched for the maneuvering trials. This discrepancy is expected to be corrected by a dynamic attentional submodel. A comparison of

†At least first-order dynamics are required by the OCM.

model-data normalized PSD ensemble statistics is shown in Figs. 1-2 for the TF ensemble. The results are in excellent agreement. A t-test performed pointwise was used to confirm the equality of the PSD means at the 95% confidence level.

TABLE 1 - COMPARISON OF TEMPORAL TRACKING ERRORS, TOW SYSTEM

	M	Azimuth SD		Elevation SD	
		Data	Model	Data	Model
TF	23	0.106(0.017)	0.119(0.018)	0.050(0.009)	0.052(0.009)
TS	19	0.057(0.005)	0.054(0.004)	0.030(0.006)	0.031(0.005)
TM	19	0.35 (0.10)	0.35 (0.07)	0.14 (0.02)	0.05 (0.01)

Individual runs produced by model and data can also be compared subjectively. Fig 3 is a comparison of a model and a data run from the TF ensemble. Fig 4 likewise is a comparison of model-vs-data trials for one of the maneuvering target passes. The "eyeball" similarity is quite impressive.

DRAGON System

The DRAGON is a shoulder-mounted system that consists of a launch tube plus sight. The front of the tube is pivoted on a support; the rear part of the tube rests on the operator's shoulder. Thus, as the operator (usually in a seated position) tracks a crossing target he must continuously move his shoulder by leaning his torso more and more to one side. There are no dynamics associated with the system per-se. The only dynamics are those associated with the operator's torso--i.e. the control "manipulator". These dynamics are approximated as

$$T(s) = \frac{s/\beta + 1}{\left(\frac{s}{\omega_n}\right)^2 + 2\,\zeta\left(\frac{s}{\omega_n}\right) + 1} \qquad (15)$$

where $\omega_n = 11 \pm 1$ rad/sec, $\beta = 3 \pm 1$ and $\zeta = .15$ for tilt in the azimuth axis.

The motor-noise for the DRAGON system is assumed to scale with commanded angle/body tilt as randomness increases greatly if one is required to track while leaning to one side. The motor-noise parameters are

$$V_u^o = [8,\ .25],\ \rho_u = .0001,\ \rho_{EA} = .000015$$

The weighting on error rate $Q_{\dot{e}} \simeq 0$.

The tracking error data for DRAGON consisted of $\simeq$ 40 passes of a crossing target at 1km with $|\dot{\theta}_T| = 10$ mr/sec. In approximately 1/2 of the runs the target moved from right to left (DR); in the other runs motion was from left to right (DL). A comparison of the temporal RMS statistics of model-vs-data is shown in Table 2. The results are excellent, but this is not a stationary

ensemble as the motor-noise covariance increases during the course of a run! Indeed, Figs 5-6 show the true nature of the tracking error ensemble for model and data. This is a more meaningful comparison than is temporal RMS error.

TABLE 2 - COMPARISON OF TEMPORAL TRACKING ERRORS, DRAGON SYSTEM

	M	Azimuth SD		Elevation SD	
		Data	Model	Data	Model
DR	21	7.35(3.3)	6.82(0.9)	2.78(1.5)	2.73(0.35)
DL	22	7.2 (2.6)	6.82(0.9)	2.35(1.4)	2.73(0.35)

Comparisons of normalized error PSD for model and data provides another yardstick for judging the effectiveness of the OCM application. As noted earlier, interpretation of these results must be made cautiously as the ensemble is non-stationary. Nevertheless, we can consider this as the "average frequency content" in the error waveforms. Figs 7-8 contain the model-data frequency comparisons for the azimuth axes. The results are excellent. A final model-data validation test is via the subjective comparison of individual tracking error time histories. Fig 9 shows a typical data run vs. a typical sample path from the OCM.

ITV System

In the ITV System a TOW mount is driven through rate command dynamics by the human using a handlebar controller. The system dynamics can be approximated by the transfer function

$$T(s) = \frac{K}{s(Ts + 1)} \qquad T = \frac{1}{\pi} \text{ sec}, \qquad K = .1$$

Since the handlebar is spring-loaded we assume that the motor-noise scales with commanded control input. The motor-noise parameters pertinent to ITV are

$$V_u^o = [.04, .01], \; \rho_u = .03, \; \rho_{EA} = .012$$

In addition it was found that a weighing $Q_e \simeq .5$ resulted in a best match between model and data PSD ensembles.

TABLE 3 - COMPARISON OF TEMPORAL TRACKING ERRORS, ITV SYSTEM

	M	Azimuth SD		Elevation SD	
		Data	Model	Data	Model
IC	28	0.112(0.035)	0.114(0.027)	0.092(0.023)	0.091(0.017)

There was data from only one ensemble for ITV, corresponding to a crossing target at 2Km with $\dot{\theta}_T \sim 1$. The target was moving towards the gunner on a 40^o angle. Table 3 gives the model-data comparisons for RMS tracking error in this stationary ensemble. A subjective comparison of a typical model-vs-data time history is given in Fig 10. The comparison of PSD ensemble statistic of model and data is provided in Figs 11-12. Again, we find excellent agreement between OCM results and the field-test data. Note that this agreement is excellent not only for the PSD mean statistics, but also for the 2nd-order statistics that give an indication of the run-to-run variability.

CONCLUSIONS

It has been demonstrated that the Optimal Control Model can be used to generate accurate facsimiles of target tracking error in various different anti-tank systems. While these results were not entirely unexpected, based on previous applications of the OCM, they are quite interesting in that comparisons have been made across several dimensions. By using the model to generate an ensemble, data and model ensembles can be studied, averaged and manipulated in similar manners, yielding similar results (at least to 2nd order statistics).

The types of systems studied were quite varied, especially with regard to their manipulator characteristics. Thus, it was necessary to adjust the motor-noise parameters in the OCM among systems.

Further application of the OCM to anti-tank tracking systems is expected to refine these results, focus on dynamic inter-axis attentional allocation, and refine techniques for parameter value identification.

ACKNOWLEDGEMENT

The author would like to acknowledge the personnel at the Army Material Systems Analysis Agency (AMSAA), in particular, Mr. Patrick O'Neil and Mr. Patrick Corcoran, for their support and assistance in supplying what seemed like endless numbers of tracking runs for the various systems studied.

REFERENCES

1. Kleinman, D.L., "A Model for Generating Human Tracking Error in Anti-tank Systems", ALPHATECH, Inc., Burlington, Mass. Interim Report on Contract DAAK11-80-C-0050, July 1981.

2. Kleinman, D.L., "Monte-Carlo Simulation of Human Operator Response", Technical Report TR77-1, Dept. of EECS, University of Connecticut, 1977.

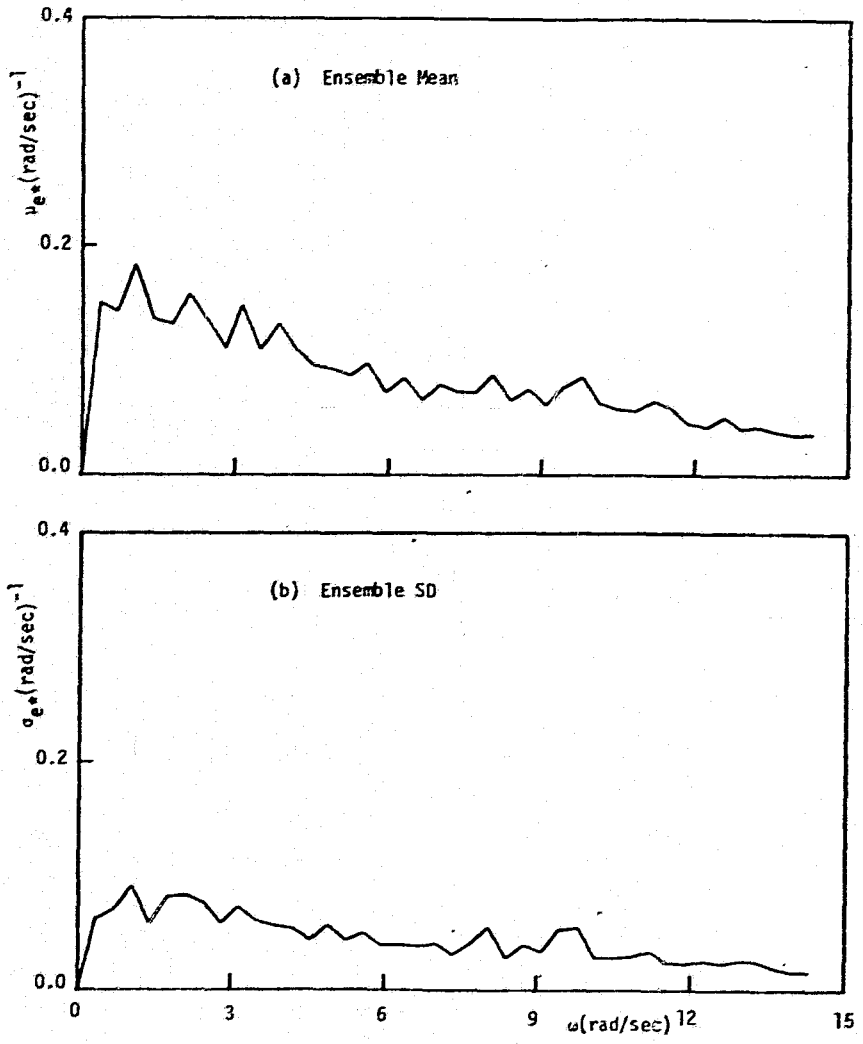


FIGURE 1- . NORMALIZED AZIMUTH ERROR PSD MEASUREMENTS, TOW CROSSING TARGET, $\dot{\theta}_T$ = 5.47 MR/SEC, M=23

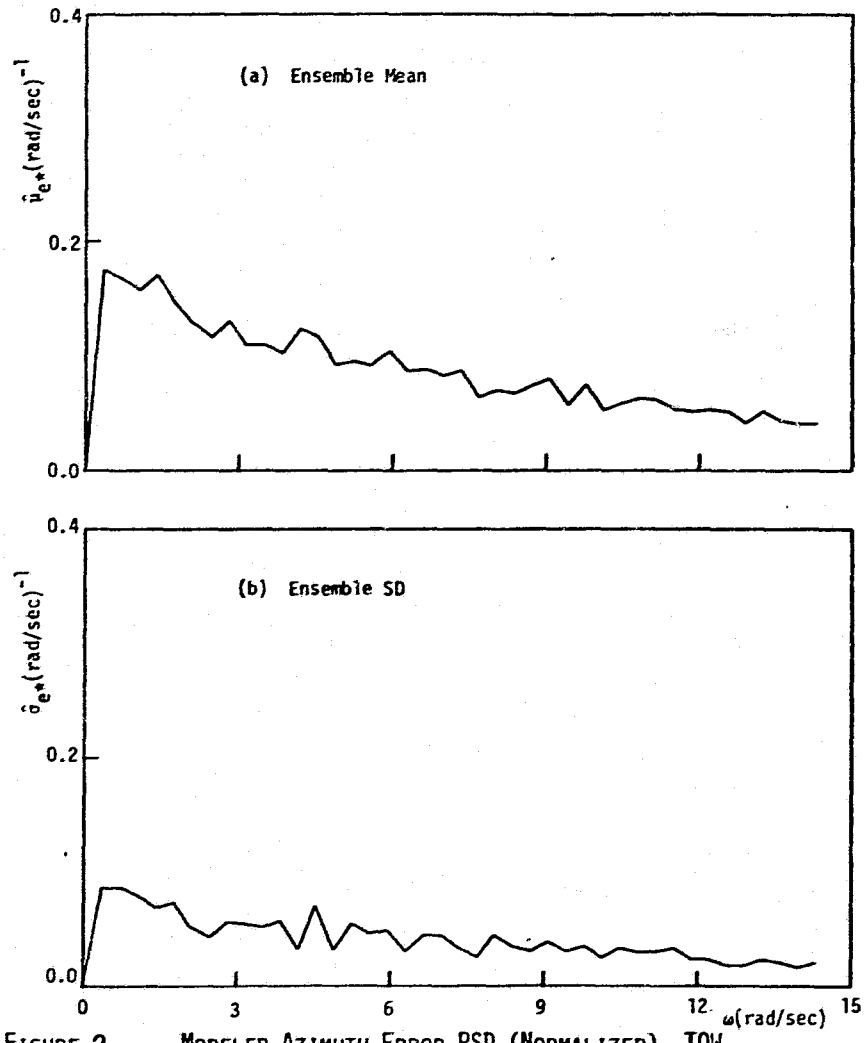


FIGURE 2- . MODELED AZIMUTH ERROR PSD (NORMALIZED), TOW CROSSING TARGET, $\dot{\theta}_T$ = 5.47 MR/SEC, M=23

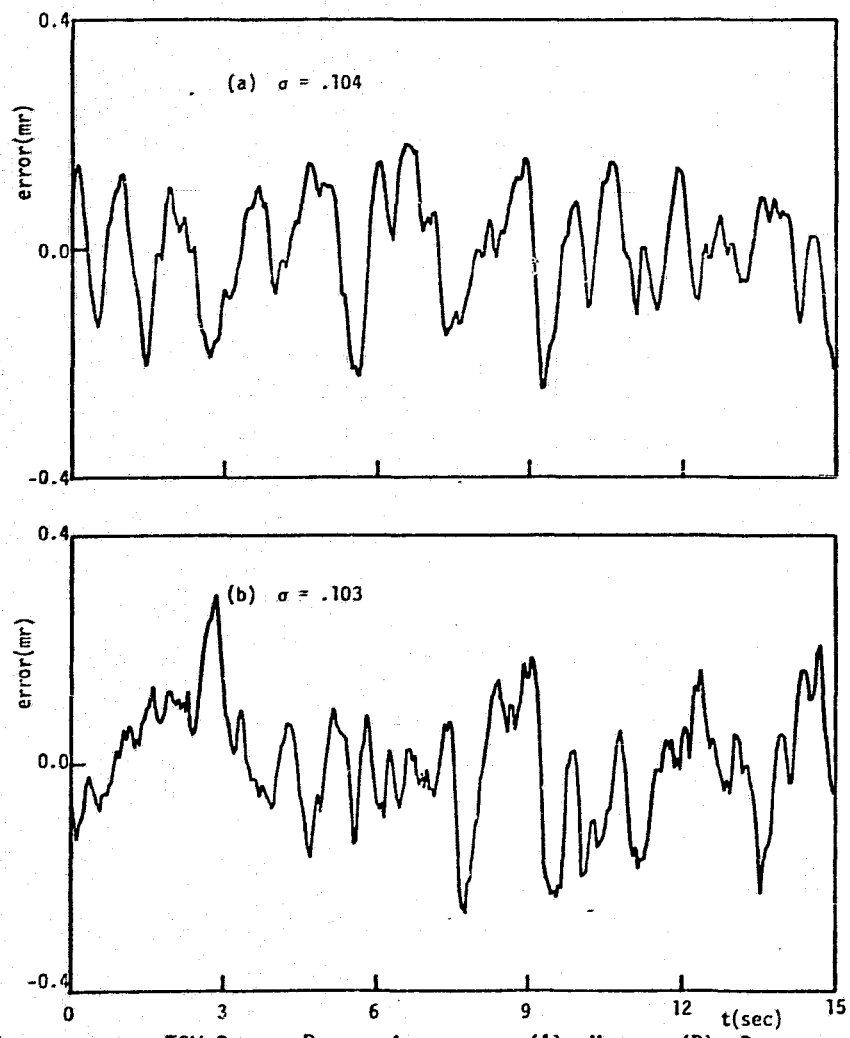


FIGURE 3- . TOW SAMPLE PATHS, AZIMUTH : (A) MODEL, (B) DATA CROSSING TARGET, $\dot{\theta}_T$ = 5.47 MR/SEC

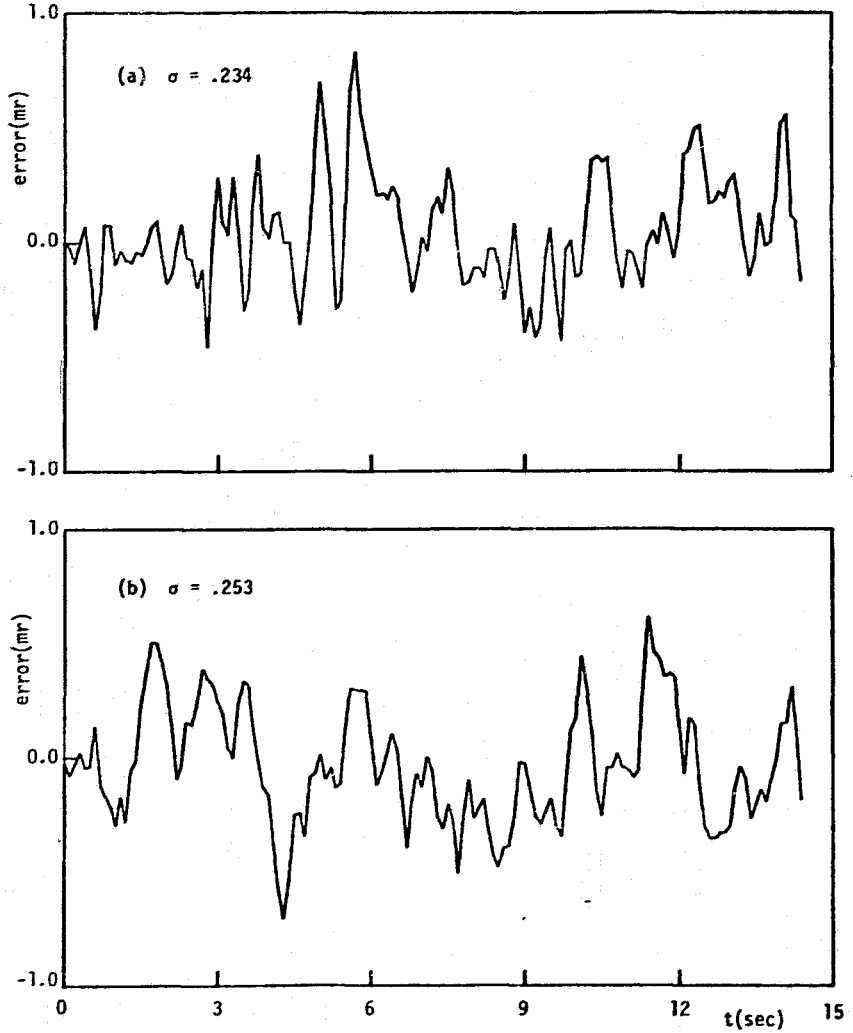


FIGURE 4- . TOW SAMPLE PATHS, AZIMUTH: (A) MODEL, (B) DATA MANEUVERING TARGET, TM12

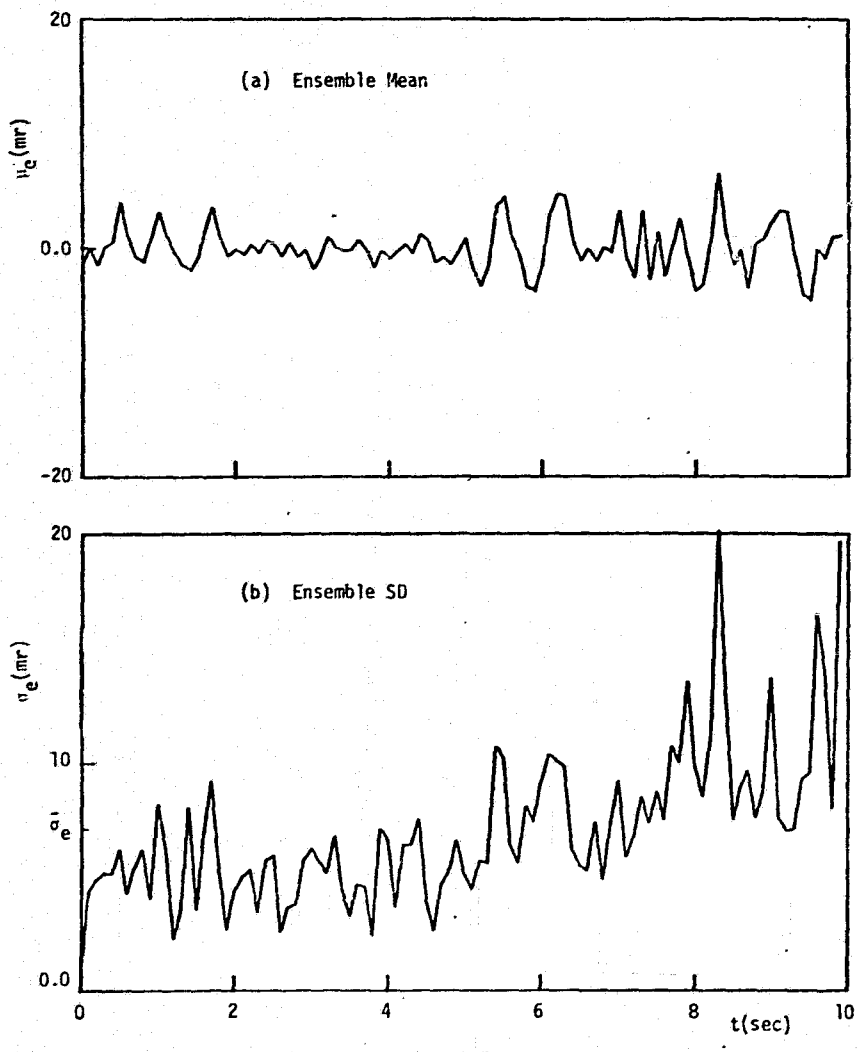


FIGURE 5- . AZIMUTH TRACKING ERROR TEMPORAL DATA, DRAGON CROSSING TARGET, $\dot{\theta}_T$ = +10 MR/SEC (R-L) M=21

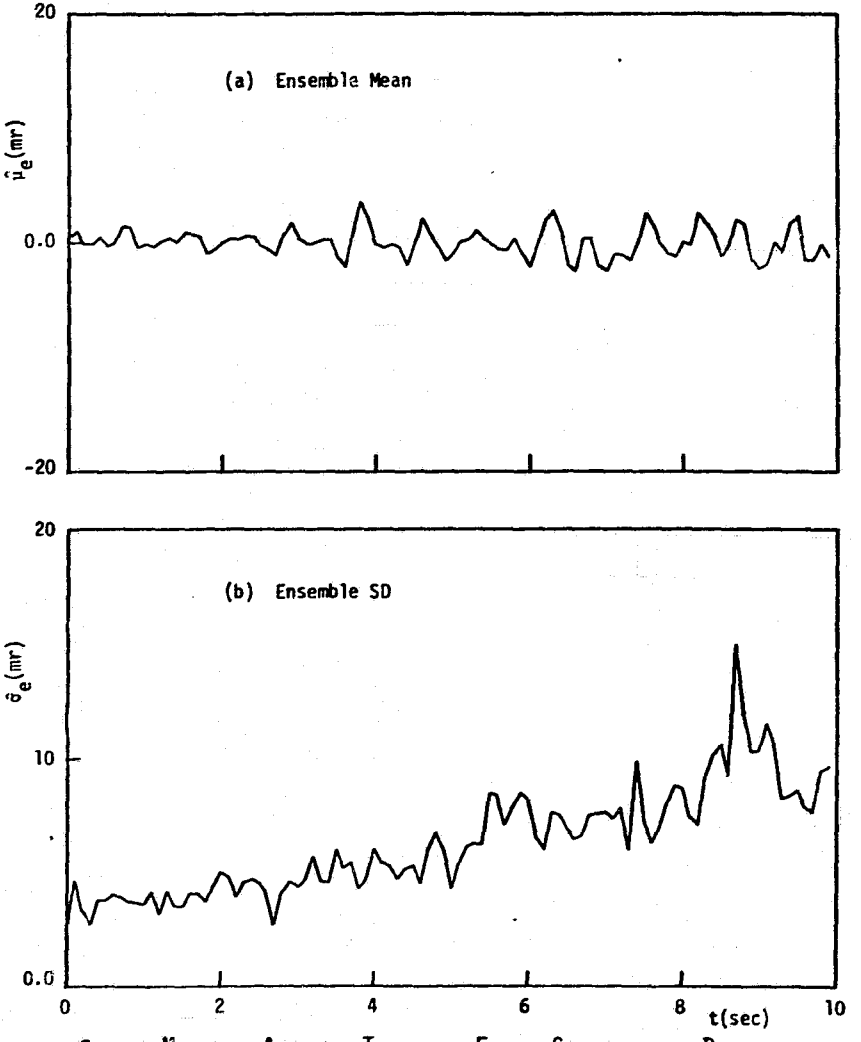


FIGURE 6- . MODELED AZIMUTH TRACKING ERROR STATISTICS, DRAGON CROSSING TARGET, $\dot{\theta}_T$ = +10 MR/SEC, M=21

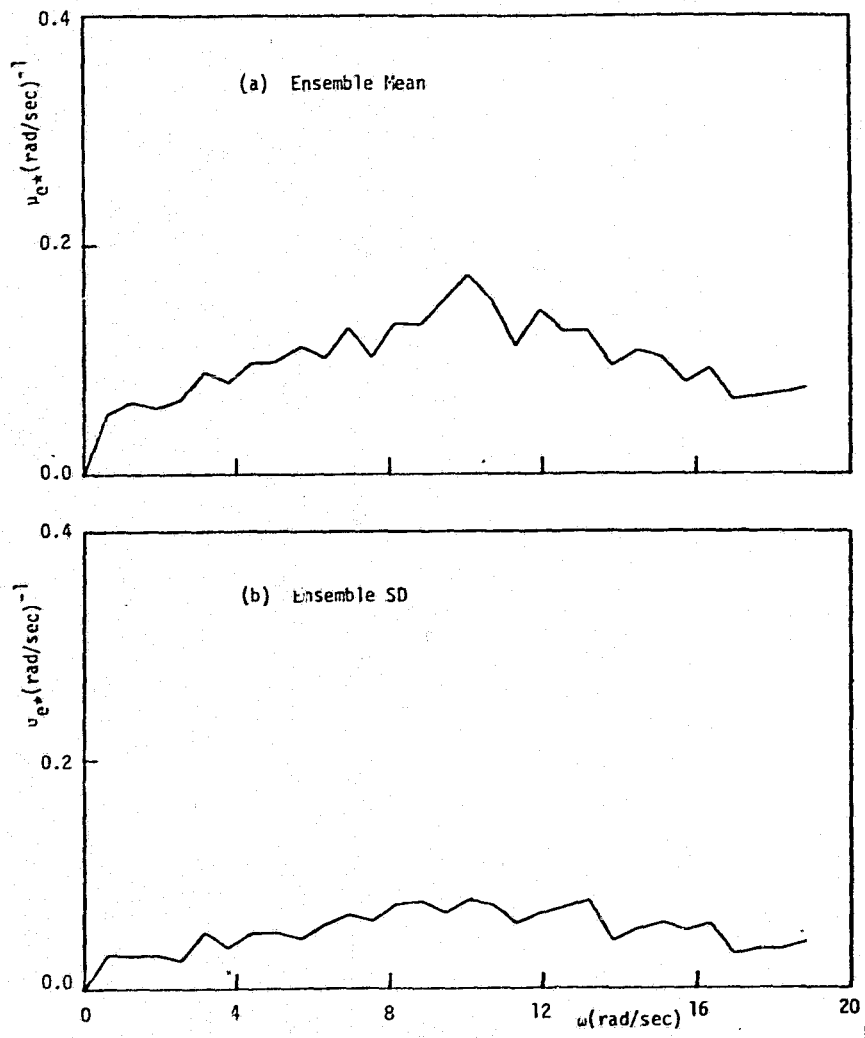


FIGURE 7- . NORMALIZED AZIMUTH ERROR PSD MEASUREMENTS, DRAGON CROSSING TARGET, $\dot{\theta}_T$ = +10 MR/SEC, (R-L) M=21

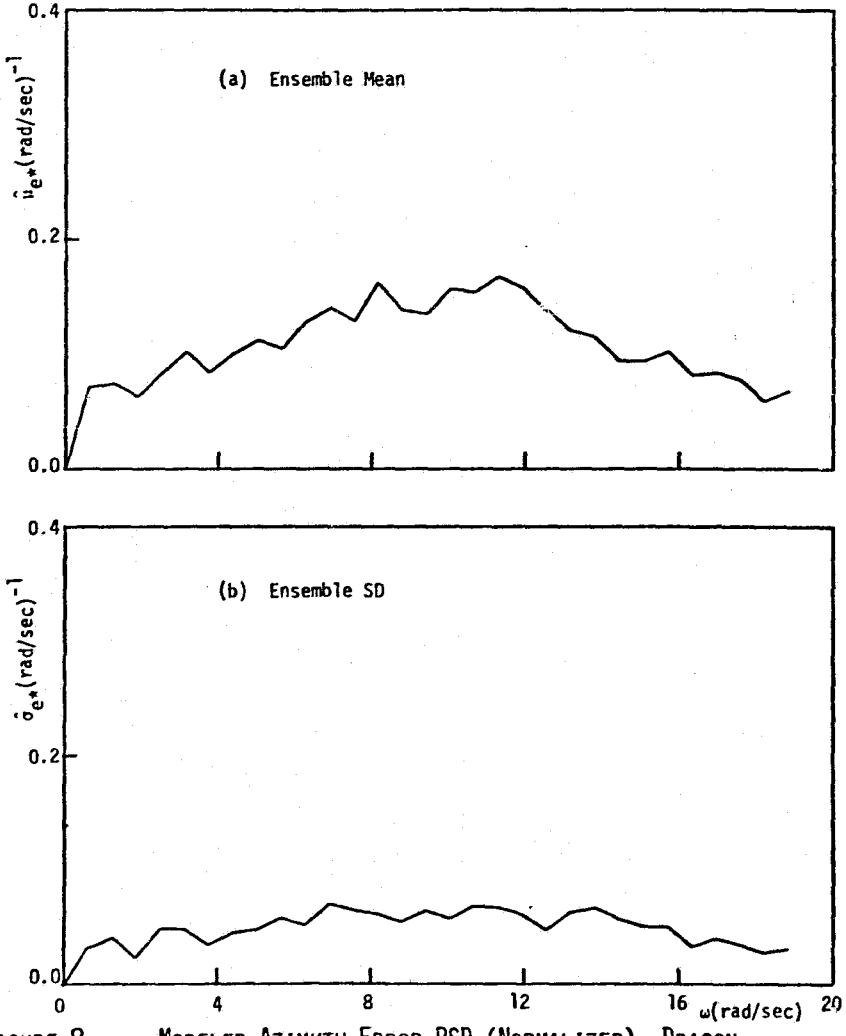


FIGURE 8- . MODELED AZIMUTH ERROR PSD (NORMALIZED), DRAGON CROSSING TARGET, $\dot{\theta}_T$ = +10 MR/SEC, M=21

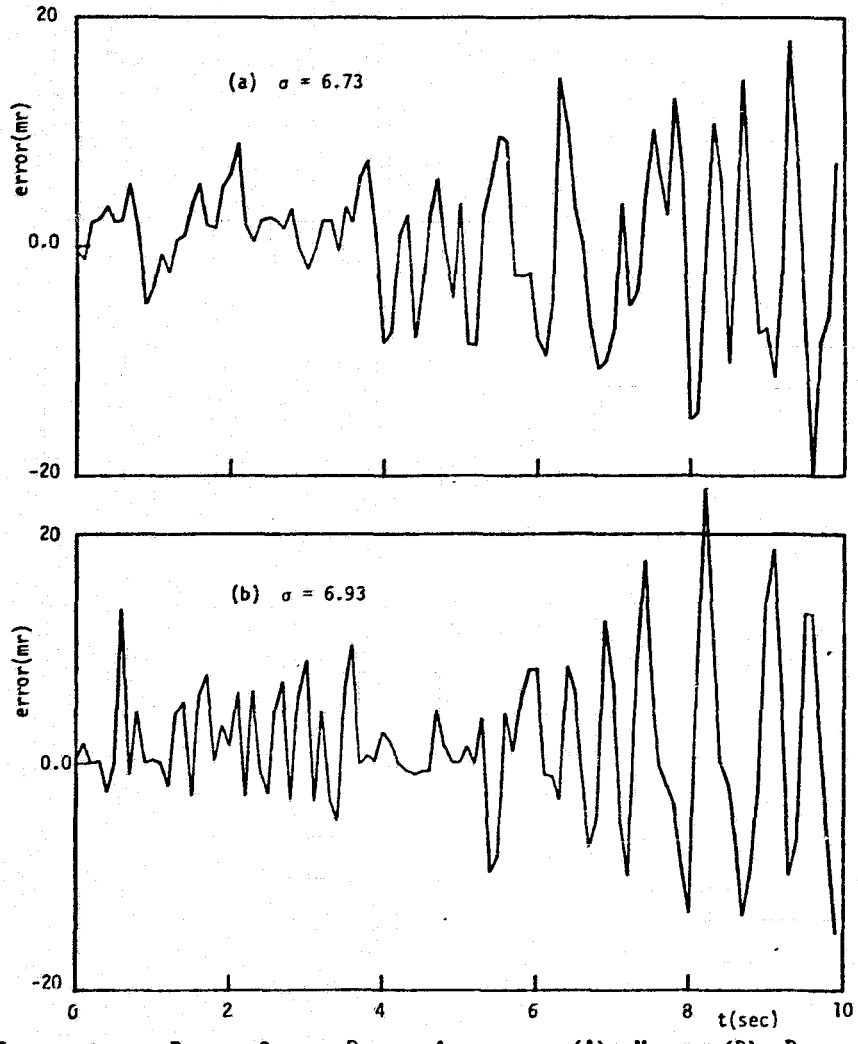


FIGURE 9- . DRAGON SAMPLE PATHS, AZIMUTH : (A) MODEL, (B) DATA CROSSING TARGET, $\dot{\theta}_T$ = +10 MR/SEC

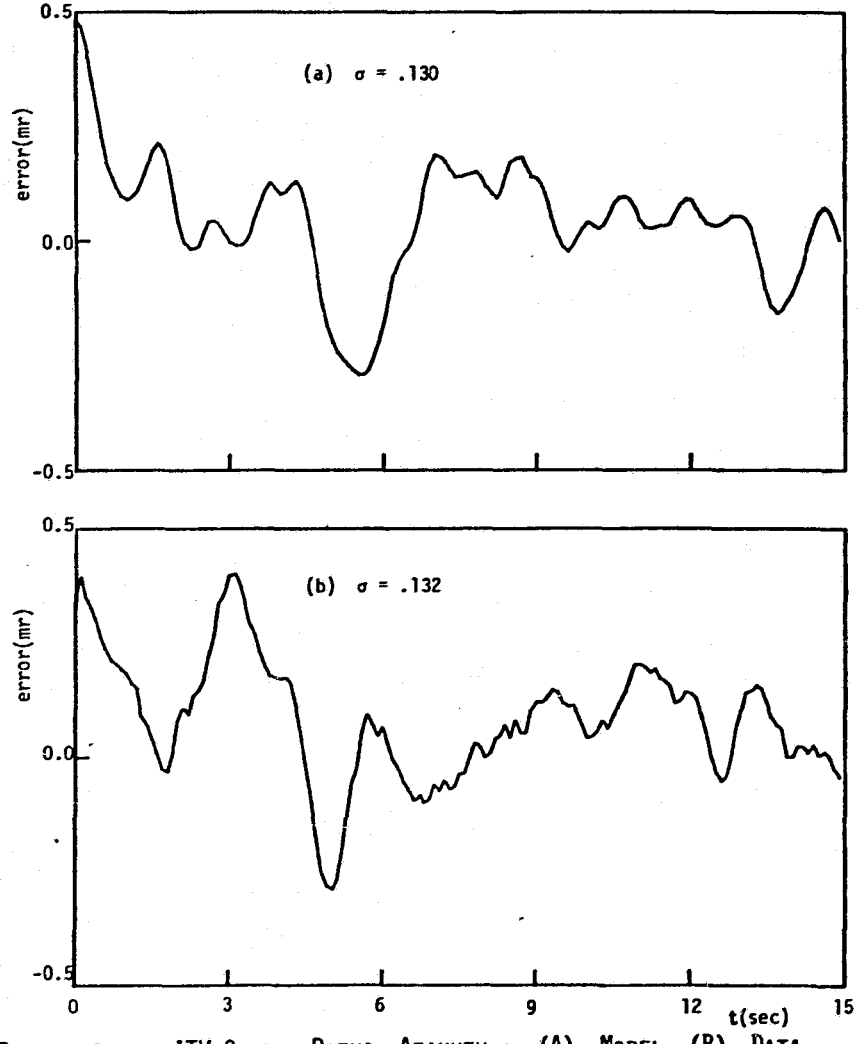


FIGURE 10-. ITV SAMPLE PATHS, AZIMUTH : (A) MODEL, (B) DATA CROSSING TARGETS, $\dot{\theta}_T$ = 1.0 MR/SEC

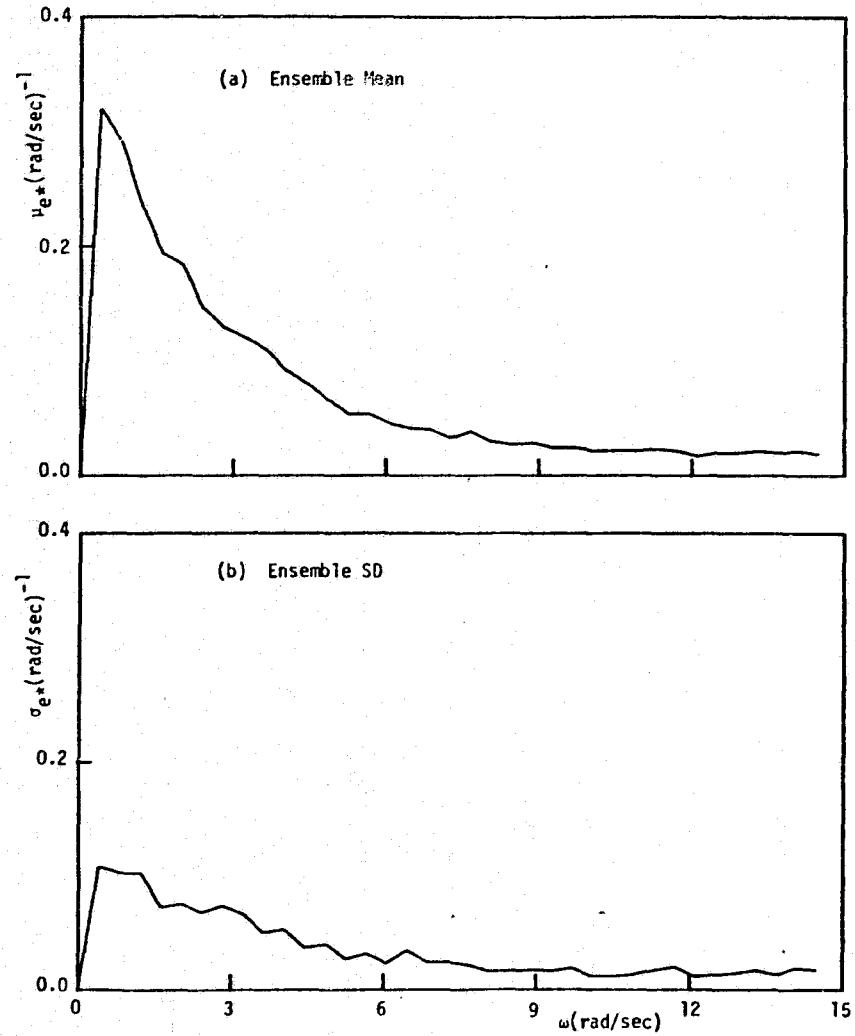


FIGURE 11-. NORMALIZED AZIMUTH ERROR PSD MEASUREMENTS, ITV CROSSING TARGET, $\dot{\theta}_T$ = 1.0 MR/SEC, M=28

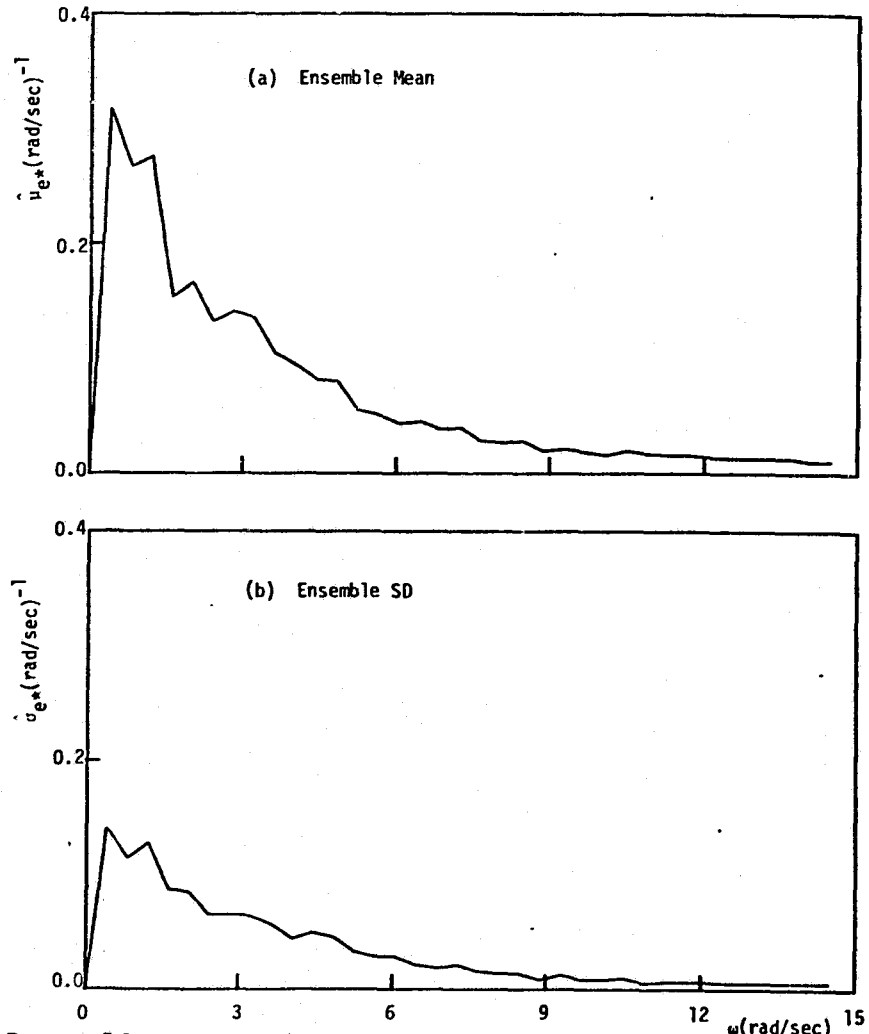


FIGURE 12-. MODELED AZIMUTH ERROR PSD (NORMALIZED), ITV CROSSING TARGET, $\dot{\theta}_T$ = 1.0 MR/SEC, M=28

ON THE INTERNAL TARGET MODEL IN A TRACKING TASK

By Alper K. Caglayan and Sheldon Baron

Bolt Beranek and Newman, Inc.

SUMMARY

In this paper, the problem of selecting a suitable internal target model for a human operator in a tracking task is investigated. The results are analyzed for a target executing a straight and level constant velocity fly-by. The internal target model is formulated in the Cartesian coordinates. The geometry of the perceived tracking error and error rate makes the formulation a nonlinear filtering problem in which the fly-by parameters are to be estimated. Although no attempt is made to match experimental data, the qualitative features of the results capture the important aspects of the empirical findings. For instance, the approach leads to a mean tracking error response which is asymmetric about crossover. The asymmetry of the mean appears to be traceable to the fundamental observability conditions arising from the formulation. As crossover is neared, the system becomes more observable and, thus, the target position and velocity estimates improve dramatically. Furthermore, the constant fly-by parameter is learned right around crossover. Given the internal model for a fly-by, this allows the gunner to estimate future position and velocity much better and, thereby reduce overshoot after crossover.

I. INTRODUCTION

The modelling of a human operator's information processing capability and control strategy in a nonstationary target tracking test has been the topic of several investigations over the years [1]-[7]. The understanding of this facet of human behavior is especially critical in manned anti-aircraft artillery (AAA) systems since the human plays a central role either as a decision maker in an automatic mode or as a feedback controller in a manual tracking mode. Therefore, the development of appropriate models mimicking the human functions of perception, decision and control in an AAA task is essential for successful manned-threat quantification predictions.

Various human operator models have been proposed for inclusion in simulations of AAA weapon systems. In the early Franklin study [1], it was assumed that the tracking interval

consists of three time invariant partitions corresponding to pre-crossover, crossover and post-crossover intervals. The authors indicate that this assumption is justifiable since the nature of the tracking test forces the operator to adopt different tracking strategies in order to prepare for high angular accelerations that will occur at crossover, to track the target at crossover, and to call off a tracking mission after crossover. The human operator model is then obtained through an impulse response matching procedure for the three intervals. While the model developed accounted for the human operator behavior up to crossover, the data matching performance was poor after crossover. Several speculations were cited for possible explanation of the post-crossover deviation such as nonsteady behavior, learning effects, nonlinear transfer behavior and cross-coupling effects.

In the Eglin study [2], the human operator model was developed using the classical control theory approach. This model contained two time varying nonlinear representations for the human operator; one for pre-crossover and the other for post-crossover tracking. Values for the human operator parameters were selected from the manual control literature and the other gain coefficients of the model were adjusted to provide a good match to tracking data. While the data matching performance of the Eglin model was satisfactory for the particular set of simulations, the Eglin model contained certain inherent limitations. For instance, the model did not predict error variance, since it did not account for the human operator variability. Moreover, this model did not explicitly account for human's adaptation to changing gun dynamics.

The use of the optimal control model for the human operator [3] resulted in predictions which were in reasonable agreement with the experimental data in the Vulcan and other studies [4]-[6]. Optimal control model provided estimates not only for the means of the variables of interest but also for the corresponding variances. In the Vulcan model, the internal model for the target trajectory was based on either a piecewise constant angular velocity (or acceleration). Furthermore, it was postulated that the human did not know the value of the incremental step change in the target's angular velocity (or acceleration). While the model predictions were improved over the previous applications, certain asymmetric and structural trends in the human response data could not be predicted.

This asymmetry is the experimental data was predicted by assuming a first order model with a variable bandwidth for the target angular velocity [7]. The bandwidth parameter was

continually updated using a specific identification scheme. In this paper, we present a different modification of the optimal control model which also predicts the asymmetry in the tracking data. Here, the target dynamics (constant velocity, straight fly-by) are exactly formulated in the rectangular coordinates and the specific geometry arising from the gunner's perception in the spherical coordinates are specified. A nonlinear filter is then employed by the gunner to estimate the target parameters. The difference between this method and the one given in [7] is that the internal target model here is dependent on the class of target maneuvers. For instance, if a different set of target trajectories are to be studied (e.g. zigzag maneuvers), then the internal model here would be changed to reflect the change in the target dynamics. In contrast, the internal target model in [7] for the zigzag maneuver would be the same as the constant fly-by case.

II. INTERNAL TARGET MODEL

The problem geometry considered here is given in Figure 1.

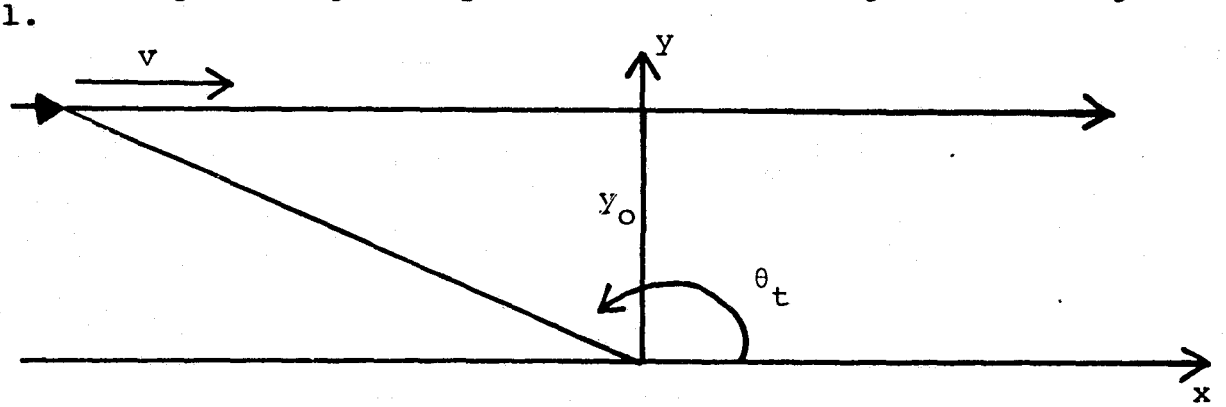


Figure 1. Problem Geometry

The target trajectory is from a class of constant velocity straight and level fly-bys. The constant target velocity, v, and the range at crossover, y_o, are unknown to the gunner. The human controls the rate of the sight angle, θ_s, in order to minimize the observed tracking error, $\theta_T-\theta_S$, where θ_T is the target azimuth. Defining the state variables by

$$x_1=\cot(\theta_T) \quad x_2=v/y_o \quad x_3=\dot{\theta}_T \quad x_4=\theta_T-\theta_S \quad x_5=u \qquad (1)$$

we get the following internal model for the target dynamics:

$$\dot{x}_1=x_2 \quad \dot{x}_2=0 \quad \dot{x}_3=2x_1x_3^2 \quad \dot{x}_4=x_3-x_5 \quad \dot{x}_5=\dot{u}+v_u \qquad (2)$$

where u is the gunner control and v_u is the operator motor noise. The gunner perceives the tracking error and derives the rate of this error so that the measurement equations become:

$$y_1=x_4+v_e \qquad y_2=x_3-x_5+v_{\dot{e}} \qquad (3)$$

We have utilized an extended Kalman filter for the gunner's estimator based on equations (2) and (3). We are postulating that the gunner knows that the target is executing a straight and level fly-by. However, he does not know the target's velocity and range. As can be seen from Figures 2 and 3, the gunner's estimate of target position (θ_T) and velocity (v/y_o) improve dramatically as crossover is neared. This behavior is expected due to the fundamental observability conditions arising from the problem geometry. The learning of the constant fly-by parameters right around crossover is the main reason for the asymmetric mean tracking response shown in Figure 4. Figures 4 and 5 are the ensemble average of 15 model runs. As can be seen from Figure 5, the standard deviation of the tracking error also captures the trends of the empirical findings in [7]. We have used nominal model parameters in these runs. The steady-state control gains were used for the linear model utilizing the current estimates for x_1 and x_3 so that

$$\dot{\hat{u}} = (10+2\hat{x}_1\hat{x}_3)\hat{x}_3+50\hat{x}_4-10\hat{u} \qquad (4)$$

In our formulation $-2x_1x_3$ corresponds to the bandwidth in [7].

III. CONCLUSIONS

Another modification in the optimal control model for an AAA gunner is presented. Although no attempt is made to match experimental data, the qualitative features of the results predict the important aspects of empirical findings such as asymmetric mean tracking error. Since the developed internal target model is dependent on the specific class of target maneuvers, the approach of this paper may be useful in modelling the decision making process of a gunner in identifying a specific target maneuver out of a possible number of target trajectory classes.

REFERENCES

1. Rosenberg, B.L. and Segal, R.: Dynamic Response of the Human Operator in Tactical Fire Control Situations. The Franklin Institute Research Lab. Report FF-C2016, April 1969.

2. Planchard, J., Barzinji, J., and Perkins, T.: Determination of Tracking Errors in an Anti-aircraft Environment. Air Force Armament Laboratory, Florida, AFATL-TR-72-165, August 1972.

3. Kleinman, D.L., Baron, S., and Levison, W.H.: A Control Theoretic Approach to Manned-Vehicle Systems Analysis. IEEE Trans. Autom. Control, Vol. AC-16, No. 6, December 1971.

4. Kleinman, D.L.: A Predictive Model of the Human Operator in the Vulcan Air Defense System. Systems Control, Inc. Final Report DAAF03-72-R-0153, November 1972.

5. Baron, S., and Levison, W.H.: Analysis and Modelling of Human Performance in AAA Tracking. Bolt Beranek and Newman Inc. Final Report F33615-72-C-1637, March 1974.

6. Phatak, A.V., and Kessler, K.M.: Modelling the Human Gunner in an AAA Task. Human Factors Journal, Vol. 19, No. 5, October 1977.

7. Kleinman, D.L., Ephrath, A.R., and Rao, P.K.: A Model for the Target Tracking Ability of a Human Operator in an AAA System. Univ. of Connecticut Final Report, AFOSR 77-3126, January 1979.

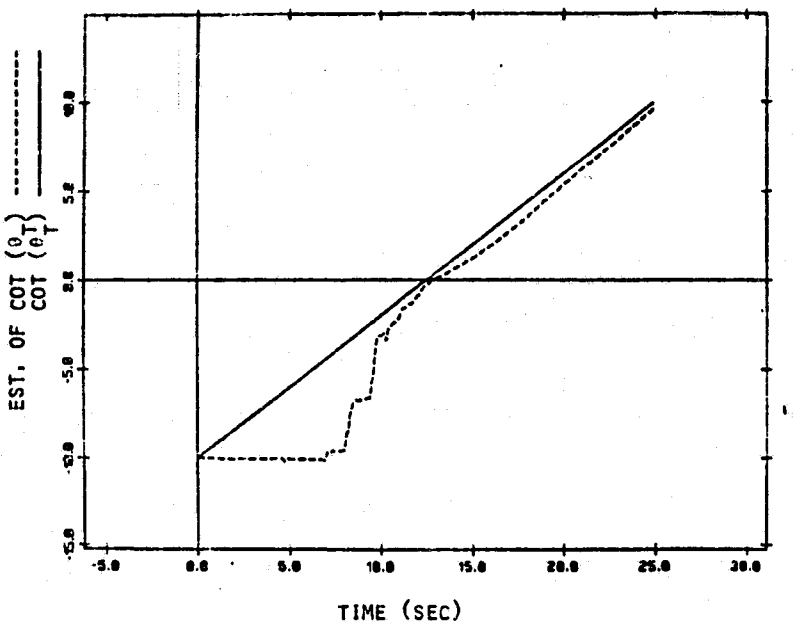


Figure 2. Estimation of Target Azimuth Angle

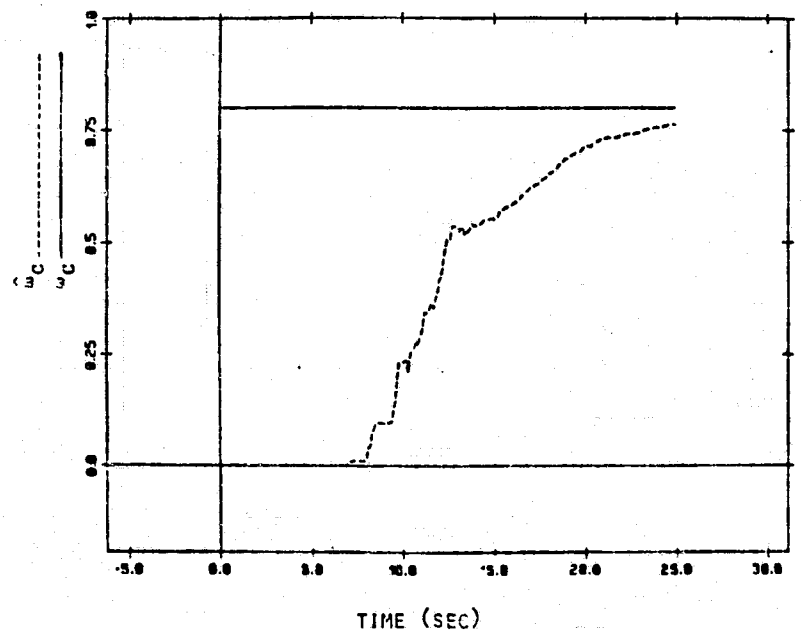


Figure 3. Estimation of Target Crossover Angular Velocity

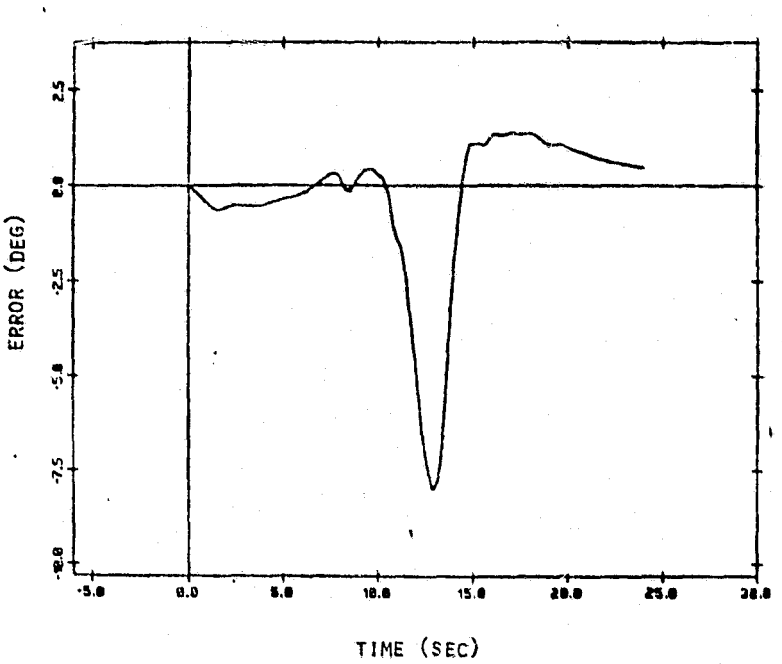


Figure 4. Ensemble Mean of Tracking Error (N=15)

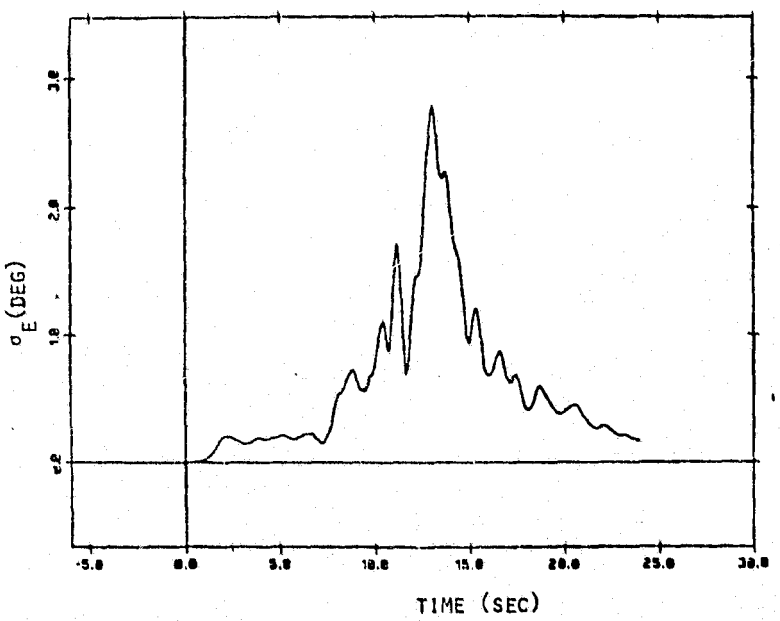


Figure 5. Ensemble Mean of Tracking Error S.D. (N=15)

D36

MULTI-AXIS MANUAL CONTROLLERS

A STATE-OF-THE ART REPORT

A. Lippay
Human Factors
Engineering

G. M. McKinnon
Manager,
Research and Development

M. L. King
Group Leader
SRMS

CAE ELECTRONICS LTD.

INTRODUCTION

Flight controls have been the subject of many studies and the interaction between pilot and controls is well documented. Typically, conventional flight controls of the joystick or wheel and column type are connected directly, or with power assist, to specific control surfaces or devices and control the movements of the vehicle. Control systems have evolved to reduce the physical effort of piloting and to generate tactile feedback signals by presenting control loading forces to the pilot's hand. This feedback is an essential factor for stability in the pilot-vehicle system and a major component of the dynamic man-machine interface.

Fly-by-wire and fly-by-computer technology have eliminated the need for direct linkage to the flight surfaces and have given rise to the concept of direct flight path control and maneuver-oriented pilot inputs. The sidearm controller was one of the first devices to emerge, reflecting the need of the cockpit designer for freedom to locate the primary flight controls away from the center of the cockpit. This need is even greater in spacecraft where the contour-seated astronaut presents special difficulties in terms of manual access to and operating envelope of primary flight controls. Furthermore, spacecraft flight demands independent control in all six degrees of freedom as well as simultaneous commands in two or more. Remote manipulators, still another newly developing technology, have similar command requirements.

PURPOSE AND SCOPE OF SEARCH

A literature search was carried out as part of a study to examine the feasibility of a six degree of freedom hand controller. A review of current work in the area of multi-axis controllers was achieved by visits to relevant research and design centres. The focus of the search was specific; however, included related areas, approaches to manual control, applications of manual controllers and selected studies of the human neuro-muscular system.

Earlier, similar efforts by the authors failed to disclose a reliable single source of information covering the field of multi-axis control devices. Furthermore, no specific titles or sections dedicated to manual control seem to exist in any of the listings and abstract journals checked.

METHODS OF SEARCH

The criteria defined for hand controller functions and related topics were transformed into descriptors recognizable by librarians and information systems. The descriptors were further adjusted as each library or service made recommendations as to the exact words to be used in defining the areas of the search.

Both direct search methods and computerized information retrieval were used. Direct search was carried out in specialized libraries, such as the Aeronautical Library of the National Research Council of Canada, Ottawa, and the technical library of the Ecole Polytechnique in Montreal. Computer searches were requested in these libraries and those of McGill and Concordia Universities of Montreal, in addition to a manual search. Abstracting services and journals were also scanned for reference and for evidence of trends or new activities in the manual control field. Manual search produced a "hit rate" of nearly 100%, computer searches 30%, abstract journals approximately 65%.

In terms of completeness, the authors are confident that the bulk of significant work in the area of multi-axis manual controllers in North America has been included, except for one important manufacturer whose disclosures are conspicuously absent. In addition, during the state of the art survey, researchers in the field were asked for references.

ANALYSIS OF RESULTS

The net results of this search turned out to be very similar to those of a 1972 search, both in volume and in content. There is a lot of interest in the general area but very few determined efforts to define a design philosophy for multi-axis controllers or to test these under representative conditions. In sharp contrast is the consistently active and well-reported research area of describing and modelling the human operator in continuous control systems. By far, the most appropriate and comprehensive activity is the ongoing research project at JPL, led by Dr.Bejczy.

Manual controller design and evaluation is usually included, as a minor task, in the development of vehicle handling characteristics. The reports dedicate much space to system aspects, but deal with the controller in a paragraph or two. In the absence of a definitive design philosophy, the best source of information would be pilot opinion and performance/preference ratings derived from full flight simulation or actual flight evaluation. However, such reports are few.

Theoretical studies of man-in-loop requirements and reports on laboratory-based experiments are more plentiful and some were found to contain useful design data. However, in general, the studies are limited to one or two degrees of freedom, and test conditions are less than representative. Frequently, results are arbitrarily extrapolated to the real world. Such conclusions must be accepted with due qualification and great caution. Developments in fields related to controller design and incidental studies show a significant increase since 1970-72.

No final answers have been found in the literature to some fundamental questions related to multi-axis manual control. Others have been investigated in part only, hence the answers are only partially valid and reliable.

Are six degrees of freedom too many to control with one arm and hand? No proof is offered, affirmative or negative. Qualitatively, the needs and requirements are well understood; the control task must become a means of accomplishing objectives, not to be a task in itself. Bejczy says the controller should be transparent to its operator, it should not in any way restrict the input commands except as dictated by a scheduled force feel system reflecting the controller system conditions to the operator.

Do mechanical properties and stick feel affect pilot/system performance? The affirmative and unequivocal statement by Kruger is supported by a large body of reports on research and development work on joysticks, grip shapes, sidearm and center stick configurations, stick forces, breakouts and gradients, ranges of motion, damping and other characteristics. The necessity and usefulness of proprioceptive feedback is accepted, but there is wide disagreement as to the nature, pattern and balance (harmony) of stick forces to be used. This, is partially due to the individual requirement of each manually controlled system and each control task.

The controller should be transparent in that the operator should feel that he is achieving the task, not merely moving a joy stick or control. This transparency is enhanced if there is a spatial correspondence between the controller and task, enabling the pilot or operator to predict the results of his manual inputs at all times with an absolute minimum of mental effort or added workload.

The isometric or force stick offers engineering advantages and reappears in the literature frequently, as a means of mechanizing the side-arm controller. Its proponents claim that since force is the principal parameter of proprioception, and since pilot comments are mostly centered on stick forces vs system response, deflection is not necessary for aircraft controls. Many of such statements are based on laboratory experiments with non-representative equipment. Some claim definite superiority for pressure (force) controls, especially with increasing task complexity. Flight tests with isometric sticks have been disappointing, but this is blamed on lack of proper understanding and application of this type of controller. A tendency to generate crosstalk between axes, poor stick feel and hand fatigue are reported most often ad drawbacks or areas of further work to be done. In summary, the superiority of isometric sticks for spacecraft application is by no means proven.

Forces appearing on the control stick, both active and reactive (resisting movement) have been the focus of interest since the early days of systematic flight control design. The principal concern is the prevention of overcontrol or overstressing the vehicle. Since stick force dynamically leads stick deflection, stick forces provide a predictive capability similar to quickening of displays and promote head-up piloting. Even passive force systems can generate a "solid feel" which spells pilot acceptance and positive stability, while negative stick stability, backlash and Coulomb friction degrade control accuracy and increase pilot workload.

How Does the Controller Fit into the Man-Machine System? The picture is by no means complete but several research efforts and trends were identified, e.g. the concept of inner/outer control loops, objective measurement of workload, and the concept of the internal model. Typically, pilot workload levels have been derived from debriefing questionnaires and pilot rating of system controllability. A more objective result can be objective result can be obtained by measuring the direct and indirect muscular effort extracted from the pilot by electromyography (EMG) and by counting the control reversals (frequency of inputs) during the time frame of a given task. "White knuckles", or unproductive nervous effort is proposed as a measure of workload stress, and EMG power spectra as a metric of local muscle fatigue, both related to controller characteristics and forces. A flight evaluation related stick sensitivity, lack of command/display harmony or cross coupling tendencies to control reversals and hence workload.

The concept of an internal normative model is relatively new, although its equivalent (body image) has been recognized in pschology and physical medicine for quite some time. The human acquires through experience and cognitive process a fast-running model of the system response he is trying to bring about. If the system fails to match this model, he either increases his workload or registers a system failure. Attempts are being made to quantify this model and relate it to tracking tasks.

Is Six-Axis Control Necessary? Whitsett says yes, prompted by MMU experience in Skylab. It may also safely be said that the control-configured aircraft and direct flight path control will eventually require command inputs in six degrees of freedom. Alternatives are tried, such as 2 x 3 degrees of freedom and foot controls. The former occupies both hands and continuous control is interrupted every time an additional manual activity is required such as adjustment of TV cameras. Foot control is generally slow, and inaccurate as shown by the Skylab experience.

Integrated controls are advocated for U.S. Army helicopters where a wounded pilot could save his crew if he could fly the helicopter with a single hand.

Is a Six-Axis Device Feasible? The literature is inconslusive. The State-of-the-Art survey found three models, and several four DOF devices. No definitive design philosophy could be found on such topics as the cascading order of axes, or the segments of the arm and hand to be used as command sources.

In terms of existing designs, several six degree of freedom controllers have been found.

a) An isometric six axis controller developed at MIT and evaluated at Marshall Space Centre, Problems have been encountered due to cross-coupling and operator fatigue.

b) A controller developed by Stark Draper Labs which includes three rotational displacement axes and three isometric force axes.

c) A six axis, floor mounted displacement unit used at Martin Marietta in conjunction with Manned Manoeuvering Unit studies.

d) An experimental six degree of freedom research tool currently in use at JPL.

e) A hard suit replica controller evaluated at NASA/AMES and at JPL.

CONCLUSIONS AND RECOMMENDATIONS

The search has been productive in terms of generating a data bank, and supporting the development of a six-axis controller model. However, gathering information in this specialized field is still a labourious process with unpredictable results. The authors appeal to the Annual Conference on Manual Control to act as a forum of information exchange, to establish descriptors and abstracting methodology to show the correct interest profile for publications and to generally promote information exchange. CAE Electronics in Montreal has a computerized data base on man-machine systems which could be expanded to the benefit of all.

THEORETICAL LINEAR APPROACH TO THE COMBINED MAN-MANIPULATOR SYSTEM IN MANUAL CONTROL OF AN AIRCRAFT

By Klaus Brauser

Messerschmitt-Boelkow-Blohm GmbH, München, FRG.

SUMMARY

A new approach to the calculation of the dynamic characacteristics of the combined man-manipulator-system in manual aircraft control has been derived from a model of the neuromuscular system similar to that described by McRuer and Magdaleno. (Ref. 1) This model combines the neuromuscular properties of man with the physical properties of the manipulator system which is introduced as pilot-manipulator model into the manual aircraft control. The assumption of man as a quasilinear and time-invariant control operator adapted to operating states - depending on the flight phases - of the control system gives rise to interesting solutions of the frequency domain transfer functions of both the man-manipulator system and the closed loop pilot-aircraft control system. It can be shown that it is necessary to introduce the complete precision pilot-manipulator model into the closed loop pilot-aircraft transfer function in order to understand the well known handling quality criteria of MIL-F-8785B/C, and to derive these criteria directly from human operator properties.

INTRODUCTION

The pioneer work on the precision pilot model presented by D.T. McRuer and his co-workers (Refs. 1...3) has become the most important step towards our understanding of the role of man in manual control. It is the combination of neurophysiology, physical dynamics, and control theory which gives the fascinating aspects of how the several complicated problems of manual control should be solved. But, there is a little gap between the theory and practical solutions, because the precision theory always turns out to be too complicated if engaged to solve such problems as the question for the best manipulator characteristics in a high performance aircraft.

In fig. 1 the precision pilot model is shown as it was presented by McRuer and Magdaleno (Ref. 2), which enlightens schematically, what is running in the neuromotor system when the pilot puts his hand (or legs) on the manipulator. For practical use however, the human engineer wants to revise this model to also parametrically based but simpler facts.

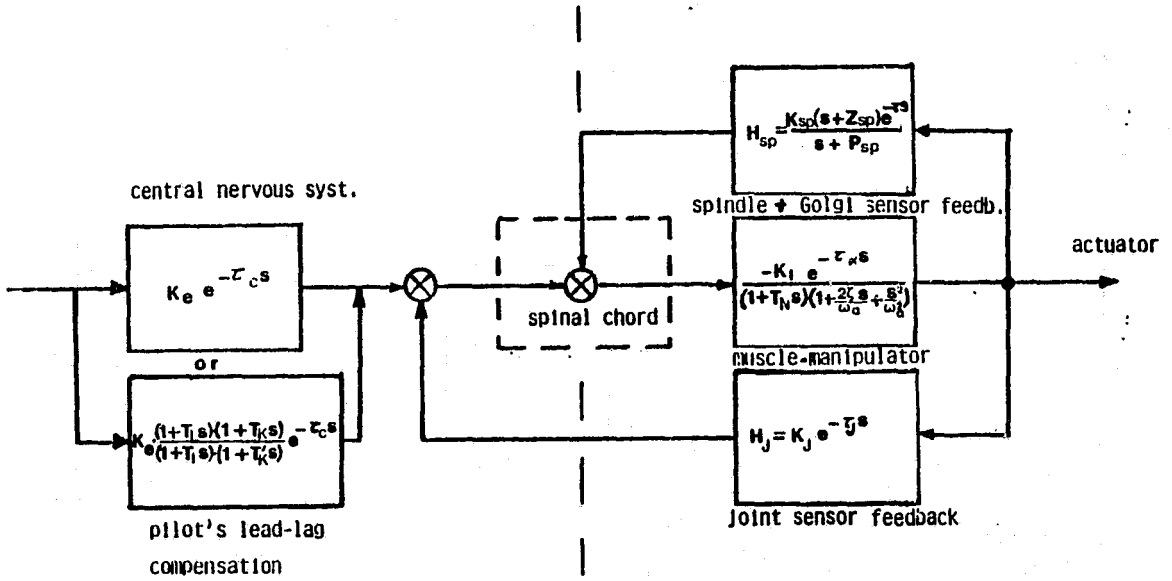


Fig. 1: Precision pilot model, after McRuer and Magdaleno (ref. 1)

If you ask the pilot about manual control qualities, he explains his wishes on

stick forces
aircraft response (as it is)
predictability of response (as he expects it is)
lead elements (after inquiry),

but he is not accustomed to consider his stick force and displacement feedback. If the pilot must consider these, the handling qualities of the system may be bad.

So for practical use we tried to reorganize the pilot model. Fig. 2 shows another pilot model less sophisticated as that of fig. 1 but more practical in use. Force and displacement feedback now feeds into the spinal chord without becoming conscious to the operator who is engaged in the compensation of system lags by lead. This model too can be fully identified by physical parameters.(Fig. 3) We now can divide the pilot model again into two parts: part A identifies the mental parameters (lead) while part B represents the neuromuscular-physical parameters of the combined man-manipulator system (lag). Now the intention of the following analysis is the formulation of the frequency domain transfer function of this quasilinear pilot model and the calculation of amplitude and phase shift of the man-manipulator combination; while further important investigations include the pilot-aircraft closed loop characteristics and their influence on handling qualities. (Ref. 5)

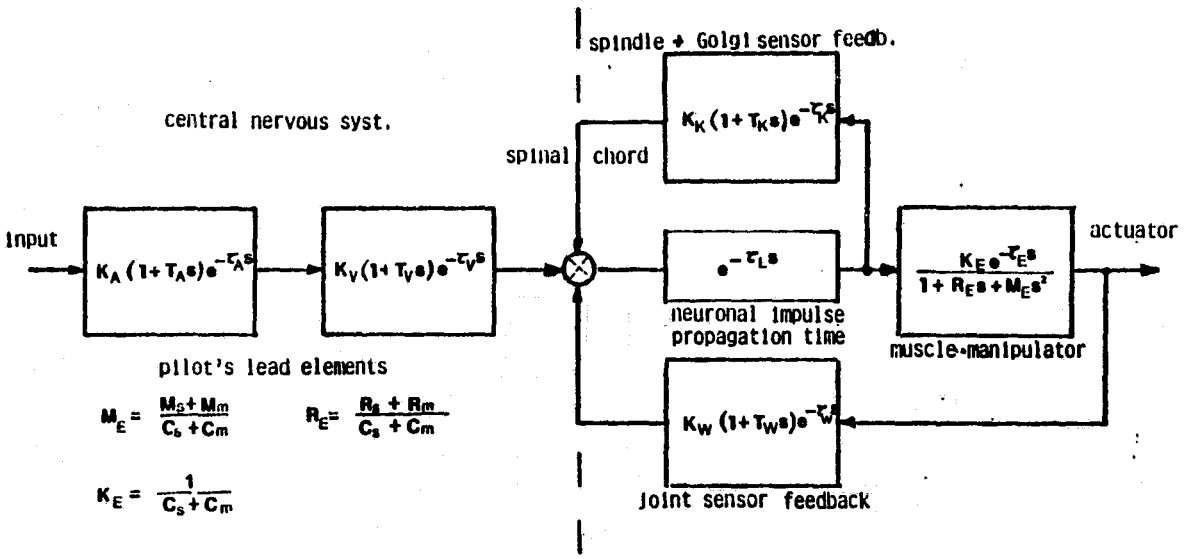


Fig. 2: Precision pilot model after P. Bubb (ref. 4), used in this report

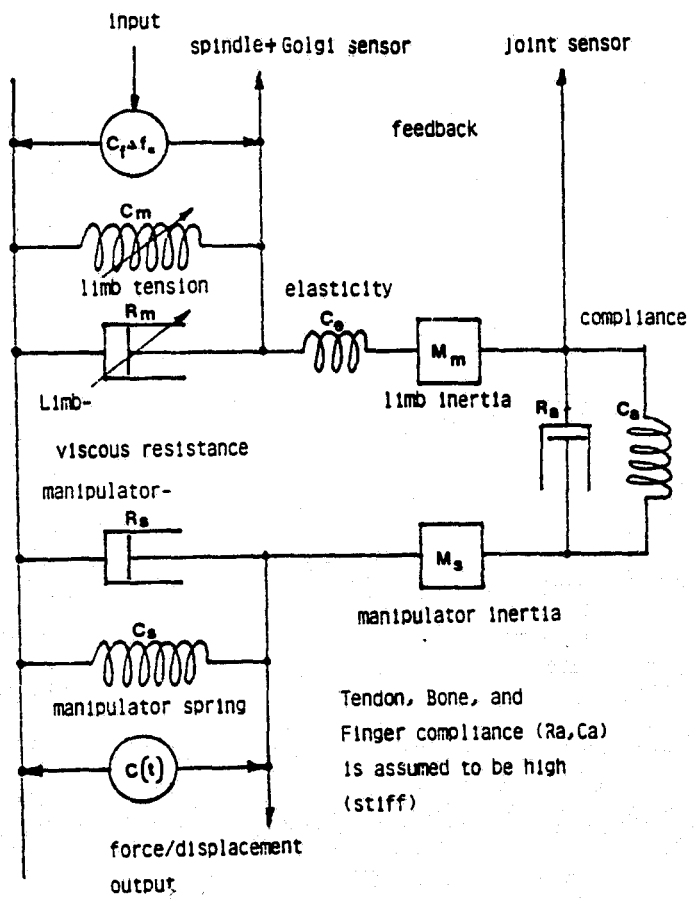


Fig. 3: Physical properties of the precision pilot model after P. Bubb (ref. 4)

ANALYSIS

Formal arrangement of the model.

The model (fig. 2) can be arranged formally alike

$$(1)\quad F_p(s) = \frac{F_A(s).F_V(s).F_E(s).F_L(s)}{1+F_K(s).F_L(s)+F_W(s).F_L(s)}\ , \text{ where}$$

$$(1.1)\quad F_A(s) = K_A.(1+T_A s)e^{-\tau_A s} \qquad \text{(information input term)}$$

$$(1.2)\quad F_V(s) = K_V.(1+T_V s)e^{-\tau_V s} \qquad \text{(information processing term)}$$

$$(1.3)\quad F_E(s) = \frac{K_E}{1+R_E s+M_E s^2} \qquad \text{(second order lag of man-manipulator)}$$

$$(1.4)\quad F_L(s) = e^{-\tau_L s} \qquad \text{(neuronal impulse delay)}$$

$$(1.5)\quad F_K(s) = K_K(1+T_K s)e^{-\tau_K s} \qquad \text{(delayed force-feedback)}$$

$$(1.6)\quad F_W(s) = K_W(1+T_W s)e^{-\tau_W s} \qquad \text{(delayed displacement feedback)}$$

Rearranging the equations (1.1)...(1.6) into eq. (1) results in

$$(2)\quad F_p(s) = A(s).B(s) =$$

$$= \frac{K_A K_V K_E (1+T_A s)(1+T_V s)e^{-(\tau_A+\tau_V+\tau_L)s}}{(1+R_E s+M_E s^2)\left\{\left[1+K_K(1+T_K s)e^{-(\tau_K+\tau_L)s}\right]+\frac{K_E K_W (1+T_W s)e^{-(\tau_W+\tau_L)s}}{1+R_E s+M_E s^2}\right\}}$$

and rewritten by the following introductions

$$(2.1)\quad K_p = K_A K_V K_E \qquad \text{(effective pilot gain)}$$

$$(2.2)\quad \tau_e = \tau_A+\tau_V+\tau_L \qquad \text{(effective pilot delay time)}$$

$$(2.3)\quad \bar{\tau} = \tau_K+\tau_L = \tau_W+\tau_L \qquad \text{(feedback delay time, since } \tau_K \approx \tau_L)$$

yields

$$(3)\quad F_p(s) = A(s)+B(s) =$$

$$= \frac{K_p(1+T_A s)(1+T_V s)e^{-\tau_e s}}{(1+R_E s+M_E s^2).\left[1+K_K(1+T_K s)e^{-\bar{\tau}s}\right]+K_E K_W(1+T_W s)e^{-\bar{\tau}s}}$$

To achieve the roots of this system, we first have to eliminate the delay terms as much as possible, which is done by multiplication of both the numerator and denominator of eq. (3) by $e^{+\bar{\tau}s}$. This results in the replacement of $e^{-\tau_e s}$ by $e^{-\bar{\tau}_e s} = e^{-(\tau_e-\bar{\tau})s}$ in the numerator, and another form of the denominator $D_p(s)$ of eq. (3) like

$$(4) \quad D_p(s) = (1+R_E s+M_E s^2)\cdot\left[e^{+\bar{\tau}s}+K_K(1+T_K s)\right]+K_E K_W(1+T_W s)$$

While the numerator has already its final root arrangement, the denominator has not. For better handling, the term $e^{+\bar{\tau}s}$ should be approximated by a Taylor series

$$(5) \quad e^{+\bar{\tau}s} = 1+\bar{\tau}s+\frac{\bar{\tau}^2 s^2}{2!}+R(s)$$

$$(5.1) \quad R(s) = \frac{\bar{\tau}^3 s^3}{n} \qquad n = 2,3,\ldots 6$$

the frequency range of which is $0 < (s=j\,) < 70$ rad/sec. with an amplitude and phase error below 20 %, for n = 2:

$$(6) \quad F_p(s) = A(s)B(s) =$$

$$= \frac{K_p(1+T_A s)(1+T_V s)e^{-\bar{\tau}_e s}}{(1+R_E s+M_E s^2)\left[1+s+\frac{\bar{\tau}^2}{z}s^2+\frac{\bar{\tau}^3}{n}s^3+K_K(1+T_K s)\right]+K_E K_W\cdot(1+T_W s)}$$

The roots of the denominator polynom are solved by the calculation of the denominator polynominal form

$$(7) \quad D_p = a_o+a_1 s+a_2 s^2+a_3 s^3+a_4 s^4+a_5 s^5$$

and by comparison of the coefficients a_i and the coefficients b_i of another, well-known polynom the roots of which meet the 5th order arrangement

$$(1+b_1 s)(1+b_2 s+b_3 s^2)(1+b_4 s+b_5 s^2)$$

The necessary arithmetics are described and discussed in ref.5. It has to be mentioned, that an exact solution of a 5th order linear arithmetic equation does not exist.

SOLUTION

The best solution of eq. (6) is the following

$$(9) \quad F_p(s) = A(s).B(s) =$$

$$= \frac{K_p(1+T_A s)(1+T_V s)e^{-\bar{\tau}_e s}}{\bar{K}(1+T_W s)\left(1+\frac{1+K_K}{\bar{K}}R_E s+\frac{1+K_K}{\bar{K}}M_E s^2\right)\left(1+\delta\frac{1+K_K}{\bar{K}}R_E s+\delta^2\frac{1+K_K}{\bar{K}}M_E s^2\right)}$$

which describes the required roots of a 5th order lag system B(s), if

$$(9.1)\quad A(s) = K_p(1+T_A s)(1+T_V s)e^{-\bar{\tau}_e s} .$$

The conditions for the solution of B(s) within eq. (9) are:

$$(10.1)\quad \bar{K} = 1+K_K+K_E K_W \qquad \text{(combined feddback gain)}$$

$$(10.2)\quad T_W = R_E = \beta\bar{\tau} \qquad \text{(neuromuscular lag time constant)}$$

$$(10.3)\quad \frac{1+K_K}{\bar{K}}, R_E = \frac{2\zeta_E}{\omega_E} \qquad \text{(2nd order lag time constant)}$$

$$(10.4)\quad \frac{1+K_K}{\bar{K}} M_E = \frac{1}{\omega_{E^2}} \qquad \text{(man-manipulator resonant frequency)}$$

$$(10.5)\quad \zeta_E = \frac{R_m+R_s}{2\sqrt{RmRs}} \geq 1.0 \qquad \text{(damping of the 2nd order lag term)}$$

$$(10.6)\quad M_E = \frac{Mm+Ms}{Cm+Cs} \qquad \text{(Inertial resistance of the manipulator)}$$

$$(10.7)\quad R_E = \frac{Rm+Rs}{Cm+Cs} \qquad \text{(viscous resistance of the man-manipulator)}$$

$$(10.8)\quad \delta = \pm\sqrt{\frac{n+2}{2nk}+\left[\frac{(Rm+Rs)^2}{2k\beta RmRs}\right]^2} - \frac{(Rm+Rs)^2}{2K\beta RmRs}$$

δ is a "high frequency weighting factor", in which β, k, and n are defined as follows

$$\beta = \frac{1+(K_K\frac{TK}{\bar{\tau}})}{1+K_K}, \quad k = \frac{1+K_K}{\bar{K}}, \quad n = 2 \text{ (see eq. 5.1.)}$$

The 5th order solution (9) degrades to a 3rd order solution, if $\delta \to 0$. A limit of δ is given by $\varepsilon < 1$: If $\delta < \varepsilon$, set $\delta = 0$, while ε is defined by

$$(10.9)\quad \varepsilon = \frac{1}{70\sqrt{k.M_E}} \qquad \text{i.e. if}$$

$$\omega'_E = \frac{1}{\sqrt{k\delta^2 M_E}} > 70 \text{ rad/sec which is the limit in } \omega \text{ de-}$$

fined by the approximation in eq. (5.1)

Numerical evaluation

While the numerical values of M_m (limb inertia), M_s (manipulator inertia), C_s (feel spring constant), and R_s (manipulator viscous resistance) can be achieved by measurement - or effectively are known - the values of C_m (limb muscular tension at operating point) or R_m (limb viscous resistance at operating point) have to be assumed in order to achieve reasonable results.

The same is done for the values of K_k and K_w which only can be defined statically as in the following way

$$K_k = \frac{F_s(\text{oper.point})}{F_s\ \text{max}} \qquad (F_s = \text{control stick force})$$

$$K_w = \frac{\delta_s(\text{operat. point})}{\delta_s\ \text{max}} \qquad (\delta_s = \text{control stick displacement})$$

while K_E might be defined as

$$K_E = \frac{F_s\ \text{max}}{F_s/n}$$

(F_s/n = stick force gradient at operating point, eg. $n = n_z$ for pitch axis)

For force-displacement-manipulators, the product $K_E K_W$ is defined to be unity at every operating point.

RESULTS

Comparison with experimental data

Some calculations of T_W, ζ_E, ω_E, and δ are made based on experimental data published in ref. 2. Most interesting data were those of ref. 2 which have been evaluated from tracking experiments with all manipulator characteristics but only the simplest controlled element (pitch axis)

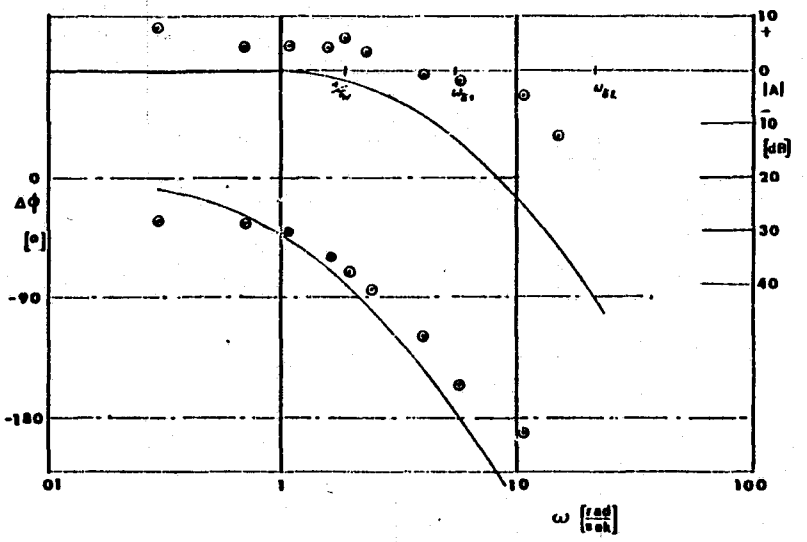
$$F_C(s) = K_C,$$

because these data are most characteristic for the man- manipulator system alone. If one assumes $K_A=K_V=1$ and $T_A=T_V \to 0$ for a controlled element without the necessity of lead compensation, which will be almost true for open loop control, the resulting transfer function is

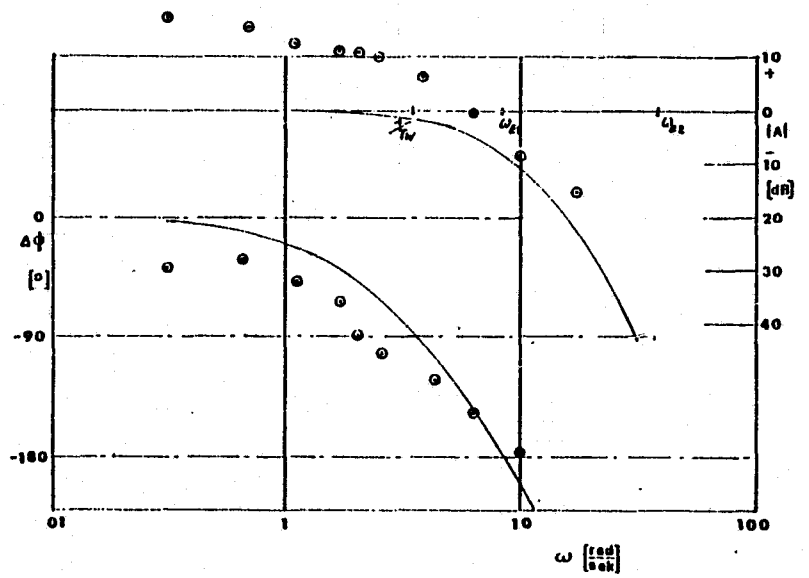
$$F(s) = B(s).K_C, \quad K_C = 1.$$

Using the data of experimental runs no. 3, 8, and 23 of ref. 2 for M_s, R_s, C_s and assuming:

run no. 3



run no. 8



run no. 23

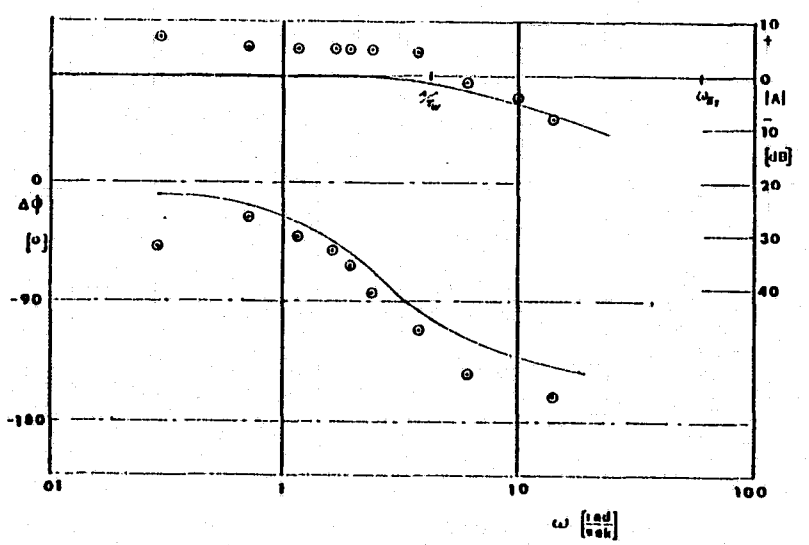


Fig. 4: Comparison between experimental data (⊙) measured by Magdaleno and McRuer (ref. 2) and theoretical data (-) - pitch axis, open loop -

M_m=0.0.0199 [kgm²] = hand-to elbow inertia for side stick

C_m=1.5 [Nm²/sec²]

$$R_m = \frac{(Mm+Ms)(Cm+Cs)}{2\zeta_s\sqrt{Ms.Cs}}$$ (see ref. 5) for ζ_s=0.7

the calculated open loop Bode characteristics are plotted against the experimental data from ref. 2 in fig. 4. As one can recognize, only the amplitude differences are quite noticeable, because the calculation was based upon K_C=1 and

K_E = 0,666 for run 3 (force-displacement stick)
0,666 " " 8 (displacement stick)
0,00913 " " 23 (force stick)

Nevertheless, the calculated plots in fig. 4 clearly show the same characteristics as the experimentally evaluated ones.

Handling quality requirements

According to the Handling quality specification MIL-F-8785-C (ref. 6) the phase angle between F_S (stick force) and δ_C (control area displacement) should not exceed-35 degrees, for level 1, related to the axis resonant frequency, ω_n. The minimum requirements for $1/T_W$ and ζ_E/ω_E are, while other phase shift between F_S and δ_C be excluded,

$$\frac{1}{T_W} \geq 3{,}5\,\omega_n$$

$$\omega_E = (3{,}5\ldots7)\,\omega_n \text{ for } \zeta_E \geq 1.0$$

Assume, the pitch axis resonant frequency ω_{nsp} = 3...5 rad/sec, these minimum requirements mean:

$$\frac{1}{T_W} = (3\ldots5).3{,}5 = 10\ldots17{,}5 \text{ rad/sec } (T_W = 0{,}057\ldots0{,}1)$$

$$\omega_E = (3\ldots5).(3{,}5\ldots7) = 10\ldots35 \text{ rad/sec}$$

As an example, we should remember the high resonant frequency of the T33 manipulator system, used by CAL during the Neal/Smith experiments, which was ω_c = 31 rad/sec (ref. 7). The low values of T_W as well as the high values of ω_E from the above assessment are realized only by light, stiff manipulators (hand controls) used aboard fighter aircraft with high resonant frequencies (ω_{nsp} = 3...10 rad/sec).

Handling quality criteria derivations

The derivation of well known handling quality criteria is possible by the introduction of the solution of eq. (9) into the closed-loop transfer function of the system pilot-aircraft. With the assumption that ω_E is high enough to satisfy the requirement

$$\omega_E \geq (3{,}5 \ldots 7) \cdot \omega_{nsp} \qquad \text{- pitch axis -}$$

the solution of eq. (9) is simplified to

$$(11) \quad F_p(s) = \frac{K_p (1+T_A s)(1+T_V s) e^{-\bar{\tau}_e s}}{\bar{K}\,(1+T_W s)} .$$

If this pilot model is used in a closed loop together with the well known pitch axis aircraft transfer function

$$(12) \quad F_{n_{sp}}(s) = \frac{\theta}{F_s} = \frac{K_c K_\theta\,(1+T_\theta s)}{s(1+T_{nsp} s + T'_{nsp} s^2)}$$

$$T_{nsp} = \frac{2\zeta_{nsp}}{\omega_{nsp}}, \qquad T'_{nsp} = \frac{1}{\omega^2_{nsp}}$$

the closed loop response is

$$(13) \quad F_G = \frac{F_p F_n}{1+F_p F_n} = \frac{\theta}{\theta_c}$$

Again, $e^{-\bar{\tau}_e s}$ has to be substituted. In this case the 1st order Padé approximation

$$(14) \quad e^{-\bar{\tau}_e s} = \frac{1-\frac{\bar{\tau}_e}{2} s}{1+\frac{\bar{\tau}_e}{2} s} = \frac{1-T_e s}{1+T_e s}$$

is the best substitution. The solution of the closed loop problem (13) can be achieved by using a method similar to that applied to eq. (6). We may suggest

$$(15) \quad F_G(s) = \frac{\theta}{\theta_c} = \frac{(1+T_A s)(1+T_V s)(1+T_\theta s)(1-T_e s)}{(1+T_W s)\left(1+\left[\frac{\bar{K}}{K_e} - \frac{T_\theta T_e}{T_n}\right] s + \frac{\bar{K}}{K_e} T_e s^2\right)(1+T_{nsp} s + T'_{nsp} s^2)}$$

with the conditions

$$(15.1) \quad T_{nsp} = T_\theta - T_e \text{ for } T_\theta > T_{nsp} \qquad \bar{K} \text{ as defined above}$$

$$T_V = \frac{T'_{nsp}}{T_{nsp}} = \frac{1}{2\zeta_{nsp}\omega_{nsp}} \qquad K_e = K_p K_c K_\theta$$

$$T_A = T_W \qquad T_e < T_e(\text{PIO})$$

These conditions show a very strong dependence between the aircraft parameters $T_{}$, ζ_{nsp}, ω_{nsp}, the most critical parameter of the man-manipulator T_W, and the lead time constants generated by the pilot himself in order to stabilize the closed loop. If level 1 proven aircraft parameters are used, the closed loop response satisfies the Neal and Smith tracking criterion (ref. 7) for good handling, if the Bandwidth of the "band-pass filter" described by eq. (15) is restricted to

$$BW = 3{,}5 \text{ rad/sec}$$
$$\log|A|(BW) = -3\text{dB}$$

The pilot conditions are then:

$$\bar{\tau}_e < 0{,}4 \text{ sec (PIO-criterion)}, \qquad \frac{\bar{K}}{K_e} \sim \frac{F_s}{n_z}$$

$$T_V + T_A < 2{,}0 \text{ sec (ref. 8)}$$

This "band-pass filter criterion" is shown by fig. 5. While T_V is the indicator for critical a/c parameters, T_A is that for critical man-manipulator parameters. Both should be as small as possible for good handling qualities.
The addition of the 2nd order term $(1+kR_E s+kM_E s^2)$ disturbes this Bandwidth conditions only, if $1/\sqrt{kM_E} \approx \omega_{nsp}$. In this case, another lead time constant T_{A1} has to be generated by the pilot in order to compensate the lag time constant kR_E, resulting in a higher order pilot lead element

$$(1+T_A s)(1+T_{A1} s)(1+T_V s),$$

the sum of which again must satisfy

$$T_A + T_{A1} + T_V < 2.0 \text{ (Arnold, ref. 8)}.$$

The higher order of the lead element surely will also degrade the pilot rating further.

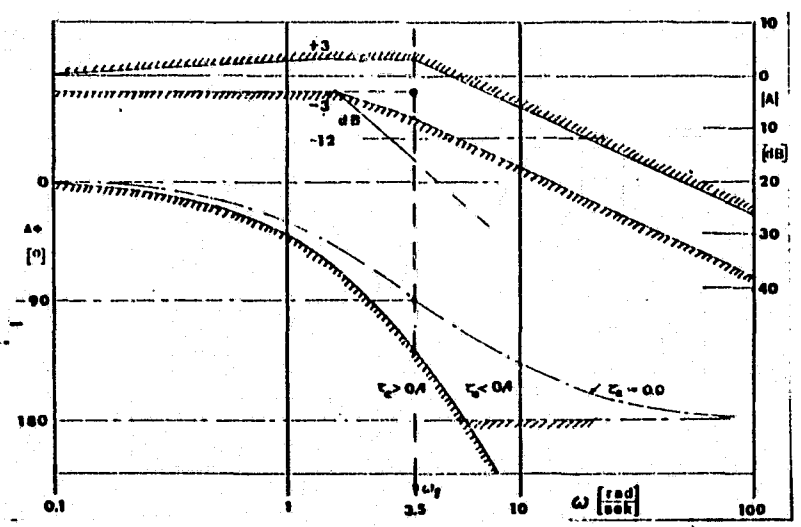


Fig. 5: Handling quality criterion proposed by Neal and Smith (ref. 7). Bode diagram of closed loop response of system pilot-aircraft

CONCLUSIONS

The formal arrangement of the full quasilinear pilot model first proposed by McRuer et. al. into both open loop and closed loop transfer functions of the manual aircraft control is a powerful means to assess aircraft handling quality criteria. The derivation of known criteria from lead term and delay term limits of the pilot has been shown clearly.

REFERENCES

1. McRuer, D.T., Krendel, E.S.: Mathematical Models of Human Pilot Behaviour
   AGARDOGRAPH AG-188 (1974)

2. Magdaleno,R.E., McRuer, DT.: Effects of manipulator Restraints on Human Operator Performance
   Technical Report AFFDL-TR-66-72 (Dec. 1966)

3. Magdaleno, R.E., McRuer, D.T.: Experimental Validation and Analytical Elaboration for models of the Pilot's Neuromuscular Subsystem in Tracking Tasks
   Report no. NASA CR-1757

4. Bubb, P.: Untersuchung über den Einfluß stochastischer Rollschwingungen auf die Steuerleistung des Menschen bei Regelstrecken unterschiedlichen Ordnungsgrades.
   Dissertation TU München, 1978. (In German)

5. Brauser, K.: Untersuchung über ein dynamisches Modell des Zusammenwirkens von Pilot und Bedienelement mit Anwendung auf die Steuerbarkeit. - in German -
   Messerschmitt-Boelkow-Blohm, TN-FE301-6/80, (July 1980)

6. Anonymous: Military Specification: Flying Qualities of piloted Airplanes
   MIL-F-8785 B (ASG) - 1969, also:
   MIL-F-8785 C /ASG) - 1980

7. Neal, T.O., Smith, R.E.: An In-Flight Investigation to Develop Control System Design Criteria for Fighter Airplanes
   AFFDL-TR-70-74, Vol. I and II (1970)

8. Arnold, J.D.: An Improved Method of predicting Aircraft Longitudinal Handling Qualities Based on the minimum Pilot Rating Concept
   Master's Thesis Nr. GGC/MA/73-1
   Air Force Flight Dynamics Lab., Wright-Patterson- AFB/Ohio
   June 1973

D 38

N82-13703

THE INFLUENCE OF SHIP MOTION ON MANUAL CONTROL SKILLS

By Peter McLeod, Christopher Poulton

MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, England

Howard Du Ross

RAE, Farnborough, Hants, England

and Wynn Lewis

WSL, Stevenage, Herts, England

SUMMARY

The effects of ship motion on a range of typical manual control skills were examined on the Warren Spring ship motion simulator driven in heave, pitch and roll by signals taken from the frigate HMS Avenger at 13 m/s (25 knots) into a force 4 wind. The motion produced a vertical r.m.s. acceleration of 0•024g, mostly between 0•1 and 0•3 Hz, with comparatively little pitch or roll. A task involving unsupported arm movements was seriously affected by the motion; a pursuit tracking task showed a reliable decrement although it was still performed reasonably well (pressure and free-moving tracking controls were affected equally by the motion); a digit keying task requiring ballistic hand movements was unaffected. There was no evidence that these effects were caused by sea-sickness.

The differing response to motion of the different tasks, from virtual destruction to no effect, suggests that a major benefit could come from an attempt to design the man/control interface on board ship around motion resistant tasks.

INTRODUCTION

Ship motion typically consists of a narrow band of high amplitude, low frequency movement with a wider band of low amplitude motion at higher frequencies superimposed on it. The degrading effects of the low amplitude, high frequency motion (i.e. vibration) on manual control skill are well known (for reviews, see Guignard and King 1972, Collins 1973, or Drennen et al. 1977) and much is known about the tendency of low frequency movement to induce nausea (e.g. O'Hanlon and McCauley 1974). But very little is known about the effects of the high amplitude, low frequency components of ship motion on manual control skills. This is presumably due, at least in part, to the high cost of building simulators to reproduce the high amplitude of the low frequency components.

The only studies of the effects of ship motion on control skills are Jex et al. (1976) and O'Hanlon et al. (1976). Jex et al. simulated the motion of

a 2000 ton surface effect ship at a variety of speeds and sea-states. Sailors spent up to 2 days in the cabin performing a range of tasks including tracking, vigilance, navigational plotting, keyboard operation and mechanical assembly. The various conditions simulated produced a range of r.m.s. vertical acceleration values between 1 and $3m/s^2$. (Note: r.m.s. acceleration is the standard deviation of the accelerations experienced during the run. For those who find it easier to appreciate acceleration magnitude in terms of g, $1m/s^2$ is almost exactly equal to 0·1g.)

The study reported here is similar to the Jex et al. study in using ship motion in three dimensions: heave, pitch and roll. However, the level of accelerations is considerably lower, in fact at a level where Jex et al. predict there will be no effects of ship motion on manual control skill. A range of manual control skills was studied: tracing (unsupported movements of the while arm); tracking, using either a pressure or a free moving control, (continuous fine hand movements with the arms supported); keyboard digit punching (ballistic movements with unsupported hands). An attempt was made to separate the effects on performance of motion itself and the effects caused by feelings of sickness induced by the motion.

METHOD

2.1. The motion

The experimental cabin was mounted on the ship motion simulator at the Department of Industry laboratory at Warren Spring, Stevenage, England (see Appendix). This was driven in heave, pitch and roll by signals recorded from the helicopter deck of the 2040 ton frigate HMS Avenger moving at 25 knots (13 m/s) into a force 4 wind. Under these conditions virtually all the motion is in heave (i.e. vertical movement): the r.m.s. accelerations in heave, pitch and roll were $0·24\ m/s^2$, $1·35°/s^2$ and $0·46°/s^2$ respectively. Given that the subject's head was about 1·7 m above the centre of rotation of the cabin the two latter figures correspond to approximately $0·045\ m/s^2$ and $0·015\ m/s^2$. These values are so low that we have only correlated performance with the vertical accelerations.

The peak to peak vertical motion was 2·5m. The average vertical r.m.s. acceleration for the hundred 7 s periods used for the tracking task was $0·31\ m/s^2$ --slightly higher than the average over the whole run. The average rate of displacement zero crossings for the whole period of the experiment corresponded to a frequency of 0·17 Hz.

Figure 1 shows the amplitude spectrum for the heave input. It can be seen that the bulk of the energy lies between 0·1 and 0·3 Hz. Figure 2 shows a typical period of 110s motion in heave. The upper trace shows the displacement signals recorded on HMS Avenger which were used to drive the experimental cabin. The superimposition of high frequency low amplitude components on the low frequency waves is clear. The centre trace shows the vertical acceleration of the cabin while being driven by the upper trace. It can be seen that a jolt, a brief period of higher than average acceleration, sometimes followed the start of an upward movement by the cabin. The average duration of the jolts was 0·28s, range 0·12 to 0·38s. The average

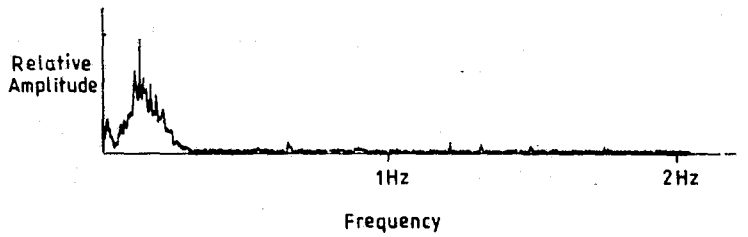


Figure 1. Amplitude versus frequency plot for the heave displacement input.

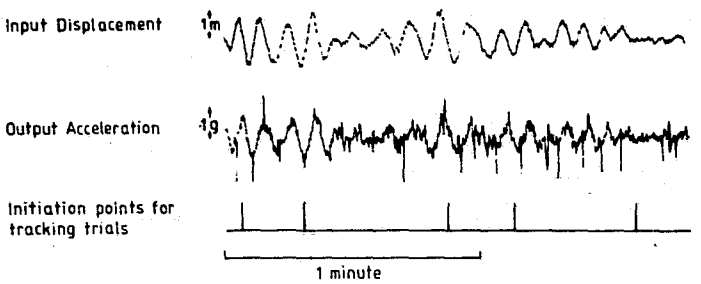


Figure 2. A 110s period of heave motion. The upper plot shows the displacement input to the cabin. The centre plot shows the vertical acceleration of the cabin. The bottom line shows the points at which tracking runs were initiated during this part of the motion.

peak acceleration was 0·16m/s^2, range 0·1 to 0·2m/s^2. This non-linearity introduced by the simulator was unfortunate but not disastrous. For the tracking task it was possible to examine the effect of the jolts by comparing performance on those trials where they occurred with those where they did not.

2.2. Motion sickness

As figure 1 shows, the motion used lies mainly between 0·1 and 0·3 Hz. This is the region which is most efficient at inducing motion sickness (O'Hanlon and McCauley 1974). (It should be noted that their data were obtained with single sinusoids: quantitative data for complex motion do not exist.) The r.m.s. acceleration value used is slightly less than that which would be expected to produce vomiting in 5% of young men after 2 hours exposure (0·33m/s^2) at the most nauseogenic frequency (0·167 Hz). However, since feelings of nausea are likely to degrade performance, it is important to try and separate these from any biomechanical effects of motion. Reason and Graybiel (1969) have reported the commonest subjective sensations which precede nausea. These sensations are a change in general well-being, dizziness, stomach awareness, headache, salivation, sweating and blurred vision. Before, during and after motion subjects rated their feelings on each of these dimensions. The subjects were given a booklet with a line 100 mm long on each page corresponding to one of the sensations with the two end points appropriately marked, e.g. 'fine' and 'awful' for the general well-being

scale. They placed a mark on each line to indicate how they felt. Changes in the position of the marks were used as an indication of which subjects felt nauseous.

2.3. The experimental cabin

A cabin measuring 2·3 m × 1·85 m with a curved roof, minimum height 1·85m was mounted on the moving platform. This was enclosed so there were no visual cues to motion for the subjects. During an experimental run the subject was strapped to a modified Sea King helicopter seat facing a console holding the CRT display for the tracking task and the LED display for the number punching task (see figure 3). Forearm restraints, the joy-stick and a numerical keyboard were attached to the deck of the console; the subject used the forearm restraints for the tracking task but not for the key punching task. An emergency button to stop the rig, a vomit bag and the booklets for subjective ratings were also attached to the console. The patterns for the tracing task were pinned to the back wall of the cabin.

Communication between subject and experimenter was via headphones; the subject was observed throughout the experiment by a closed-circuit TV camera mounted over the console.

2.4. The tasks

2.4.1. Tracing task. The subjects stood upright and tried to trace along a variety of patterns drawn on a sheet of paper pinned to the wall at shoulder height. The subjects stood at approximately arm's length from the wall and were not allowed to steady themselves by holding onto the cabin or to try and support their writing arm on the wall. They performed a set of six tracings twice over on each occasion. Measures taken were accuracy and

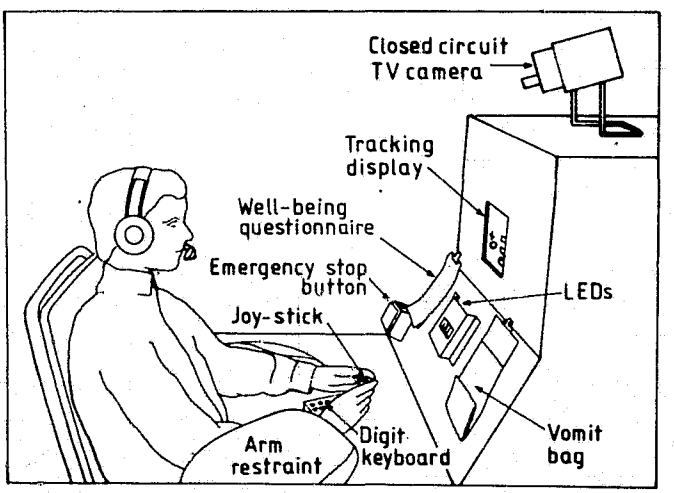


Figure 3. The console in the experimental cabin.

time to complete each set of six. The accuracy was measured by sampling the perpendicular distance from the tracing to the line the subjects were trying to follow at 31 points distributed across the six patterns. Figure 4 shows four attempts to follow the tracing patterns, the upper two static and the lower two made under motion.

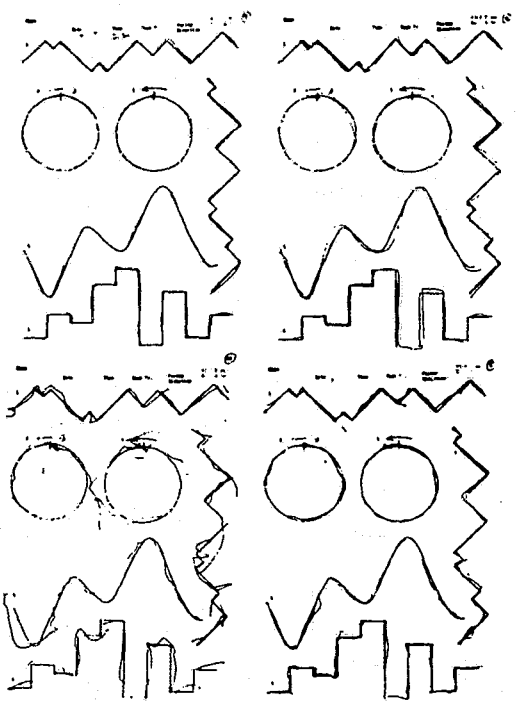


Figure 4. Four attempts to follow the tracing pattern. The upper two were done with the cabin stationary. The lower two while it was under motion. They show the approximate range from best to worst under both conditions.

2.4.2. Tracking task. This was a pursuit task, with each trial lasting 7s. The seated subjects face a 100 mm × 80mm screen at a viewing distance of about 600mm. Their forearms were supported by arm restraints. Each trial was preceded by the word READY on the screen for 1s. Then the target, a circle radius 2·5mm, and a cross (arm length 5 mm) which was controlled by the subject appeared on the screen. The cross and circle started in random positions with the proviso that the circle was inside a central area measuring 50 mm by 40 mm and the cross was outside this area. Throughout each trial the circle continued to move at random within this inner area. The algorithm used to control the movement of the circle was that every 300ms its vertical and horizontal velocities were changed independently by a random amount with a random sign up to a maximum in either direction of 0·75cm/s. The subject's task was to place the cross inside the circle as quickly as possible and keep it there for the remainder of the trial.

There were two groups of subjects, four men and one woman in each. One group tracked with a pressure control (i.e. a joy-stick which does not move,

but which gives an output proportional to the force applied to it); the other tracked with a spring-centered free-moving joy-stick, the output of which was proportional to its displacement from the central point. The relation of the position of the cross on the screen to the output of the joy-stick was

$$P_c = P_j + 3\int P_j dt + \int\int P_j dt$$

where P_c is the position of the cross and P_j the output from the joy-stick. The control, therefore, was basically a velocity control, with small components of position and acceleration.

The performance measures taken were the time to acquire the target and the modulus mean error after acquisition. "Acquisition' was defined as holding the centre of the cross within 5 mm of the centre of the circle continuously for 1 s.

The tape driving the cabin lasted for 22 min. During this time there were 50 tracking trials occurring at fixed positions at intervals of 20-30s. During an experimental run of 100 trials the subject experienced the tape twice through separated by a static period of about 25 s. The cabin went through the same motion on any particular trial for every subject.

2.4.3. Digit keying task. The subjects were presented with a series of 50 four digit numbers which they entered on a conventional calculator keyboard. Their arms rested on a horizontal surface but were not restrained. The keys were 9 mm square with a vertical and horizontal inter-key spacing of 6•5mm. In pre-motion training the numbers were spoken aloud by the experimenter. Under motion they appeared on the LEDs in front of the subject (see figure 3). In both cases the subjects were required to say the number aloud and then enter it as a group of four keystrokes. They were instructed to go as fast as was compatible with error free performance. Entry time of each key stroke and any errors were recorded.

2.5. Experimental design

2.5.1. Pre-motion practice. The subjects practiced the three tasks over a period of 3 months before going to the motion simulator. On eight separate days they performed a block of 20 tracking trials. They performed the tracing task twice and also had two blocks of 50 numbers on the digit keying task on each of two days.

2.5.2. Motion. Each subject performed the tasks under motion on two consecutive days. The complete session including stationary control trials, filling in well-being questionnaires and taking transmissibility measures lasted about 2•5 hours. The experimental design for the motion sessions is shown in table 1.

2.6. Subjects

The subjects were eight men and two women from the Applied Psychology Unit staff. They all claimed not to be prone to sea-sickness. They were right-handed, and their ages ranged from 23 to 60 years. For the tracking task they were divided into two groups of four men and a women, one group

using the pressure control and the other the free-moving control.

Table 1. The order of tasks during a motion session.

Cabin	Task	Approximate duration (min)
Stationary	20 tracking trials	15
	Well-being ratings	2
	2 tracings	2
Moving	100 tracking trials	50
	Well-being ratings	2
	2 tracings	2
Stationary	15 tracking trials	10
	Well-being ratings	2
	2 tracings	2
	5 min break outside cabin	
Motion	Transmissibility measures taken from subject	15
	Well-being ratings	2
	Key-punching task	20
	Well-being ratings	2
	Transmissibility measures taken from subject	10
Stationary	Well-being ratings	2

RESULTS

3.1. Nausea

None of the subjects actually vomited. However, there was a small but reliable drop in the feeling of well-being. Comparing the estimate made immediately prior to motion with that made at the end of the first motion session (see table 1) gives a drop of 9% in the scale going from 'Fine' to 'Awful, about to vomit'. Pooling across days and subjects this is reliable, $p<0{\cdot}05$, Wilcoxon test, 2-tail. None of the individual indices (dizziness, sweating, headache, stomach awareness, salivation or blurred vision) showed a reliable change when pooled across subjects and days. At the end of the second motion session the position was very similar. Compared to the pre-motion ratings, 'well-being' pooled across subjects and days showed a reliable 7% decline ($p<0{\cdot}05$, Wilcoxon test, 2-tail). But none of the other individual indices showed a reliable change. Table 2 shows the detailed ratings.

3.2. Tracking task

Figure 5 shows the basic performance data for the tracking task. The two left hand graphs show the performance of the group who tracked with the pressure control; the right-hand side shows the performance of the group with the free-moving control. The two upper graphs show the acquisition time (in seconds); the two lower graphs show the average modulus mean error after acquisition (in millimetres). (It should be noted that the minimum possible acquisition time is greater than zero, approximately 2s, but the minimum possible error is zero.) In each graph the average performance on the 20

Table 2. Mean subjective judgments between 100=fine and 0=awful.

Judgment	Pre-motion	Motion	Change for worse	Post-motion
Well-being	91	82	9*	88
Dizziness	90	89	1	89
Sweaty	64	55	9	61
Headache	89	86	3	88
Stomach awareness	90	83	7	88
Salivation	43	48	−5	46
Blurred vision	94	92	2	94

*$p<0.05$.

pre-motion trials and the 15 post-motion trials is shown separately, before and after the 100 motion trials. The 100 motion trials have been broken down into ten consecutive groups of ten.

The main effect, that tracking is worse under motion, is immediately obvious. Taking the mean of the pre-motion and post-motion trials as a control, every subject in both groups takes longer to acquire the target under motion ($p<0.01$, sign test). (Pressure control--acquisition time: control = 2·8 s, motion = 3·1s; error: control = 1·3 mm, motion = 1·8 mm; free-moving control--acquisition time: control = 3·0 s, motion = 3·2s; error: control = 1·9 mm, motion = 2·2 mm.) If this effect were caused by

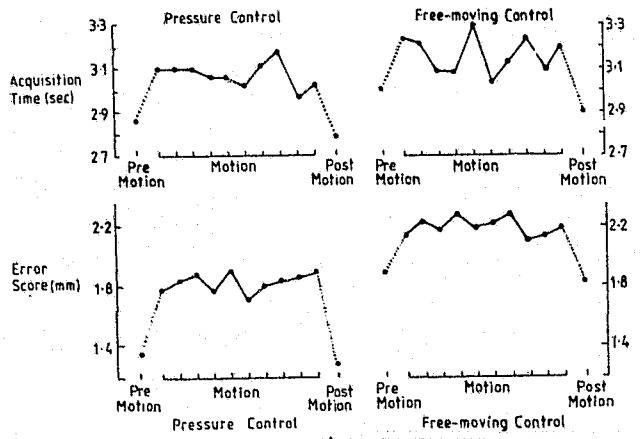


Figure 5. The results of the tracking experiment. The two left-hand graphs show the data from the pressure control. The free-moving results are on the right. The two upper graphs show the acquisition time; the two lower ones show the error score.

onset of nausea the size of the motion decrement would increase over the 100 trials (about 50 min of motion). Figure 5 shows that this is not the case. None of the four performance indices show a reliable correlation with time on task (the largest of the four correlations is 0·06). In other words the decrement caused by motion appears as soon as the cabin starts moving and is no worse after 50 min motion. Therefore we can say with confidence that there is a degradation in tracking performance caused by the biomechanical effects of this relatively small degree of motion even for subjects who are strapped to a chair and whose forearms are restrained.

It was mentioned in §2.1 that the simulator introduced some jolts, brief periods of relatively high acceleration, into the movement. It is possible to see whether the tracking decrements were caused by the jolts rather than the normal motion by comparing the performance on trials where there was a jolt with that on trials where there was not.

For both pressure and free-moving controls the presence of a jolt after acquisition made no difference to the error score. (Pressure control: error--with jolt = 1•8 mm, no jolt = 1•8 mm; free-moving control: error--with jolt = 2•2 mm, no jolt = 2•2 mm.) Clearly the motion decrement in the error score cannot be attributed to the jolts.

The presence of a jolt did produce slower acquisition by the free moving group (mean acquisition time with jolt = 3•3s, no jolt = 3•1s; $p>0•02$, Mann-Whitmey test). There was an effect in the same direction for the pressure group which failed to reach significance (mean acquisition time with jolt = 3.1s, no jolt - 3.0s; $p=0.125$, Mann-Whitney test). Part of the reduction in acquisition time under motion is caused by the jolts rather than the real ship motion. However a comparison of acquisition time on no jolt trials with the pre- and post-motion control trials shows that every subject in the free-moving group was slower on the jolt-free motion trials. Therefore, as with the error measure, it is clear that the motion does produce a reliable, if small, change in the acquisition time.

It is possible to examine the extent to which the two controls are affected by the roughness of the 'sea' by correlating the average group performance on each no-jolt trial with the mean modulus vertical acceleration on each trial. The range of mean modulus acceleration experienced ranged from 0.03 to 0.45 m/s^2. Both performance measures for both controls show positive correlations as would be expected but they were not particularly large, nor were the differences between the correlations for the two controls reliably different. (Free moving: acquisition time, $r_s=0.48, p<0.01$; error, $r_s=0.28$, $p<0.1$; pressure: acquisition time, $r_s=0.39, p<0.025$; error, $r_s=0.25, p<0.1$.) There seems to be little difference in the effects of roughness on these two controls although the rather small range of 'roughness' examined should be noted.

There seems to be little ground for deciding that either control is superior under motion. They show a similar response to both jolts and roughness. The pressure control is superior on both performance indices under motion but the same is true of the non-motion conditions. This may reflect a difference in tracking ability between the rather small groups, or a genuine superiority of the pressure control for this particular combination of task and control law.

3.3. Tracing task

The tracing patterns produced under motion were compared with those produced immediately before and after motion. The increase in error was large and shown by every subject. (Mean error: no motion=0.7mm; motion=1.9mm, $p<0.01$, sign test.) There was a small increase in the time to complete each tracing under motion, but pooled across days and subjects this was not reliable. (Mean time to completion: control=46.8s; motion=48.5s; $p=0.2$, Wilcoxon test,

2-tail.)

Figure 4 shows four tracings; the upper two were produced with the cabin stationary and the lower two under motion. The lower two demonstrate the range of tracings produced under motion from the very bad (which is typical of about half the tracings produced under motion) to one which is as accurate as the worst tracing produced when the cabin was stationary.

3.4. Digit keying task

The subjects keyed in a string of 50 four digit numbers, twice under motion and twice static. The data given for each condition are for the middle 20 four digit numbers on each of the days, pooled across the two days. The standard deviations are the means of the standard deviations for the individual subjects.

Time to enter a four digit number (static)=1069ms(S.D. 218ms)
Time to enter a four digit number (motion)=1102,s(S.D. 249ms)

The difference in mean keying time is due to chance. Half the subjects are faster under motion, half are slower. The increase in variability approaches reliability ($p=0.075$, Wilcoson test, 2-tail). There was a small increase in errors (0.5 to 1.0%) which approaches reliability across subjects ($p=0.07$, Wilcoxon test, 2-tail).

CONCLUSIONS

We have examined three manual control tasks requiring movement of the unsupported arms, continuous fine movement with restrained arms or ballistic hand movement with unsupported arms. The extent of degradation in these tasks caused by a comparatively mild ship motion is very different.

The tracing task, involving continuous whole arm movement was very seriously affected. The average error increased by a factor of three and many of the individual records were so bad that without the target tracing visible it would be difficult to guess what the subjects' intended drawing had been. The tracking task, which involved continuous fine movements of the supported arms, was reliably worse under motion, but performed with reasonable conpetence. The average error and the time to acquire the target increased by about 20%. The digit keying task, requiring a group of four pre-programmed ballistic movements, was virtually unaffected.

The changes in performance were not primarily due to nausea. Firstly, the motion was below the threshold at which 5% of people vomit after an exposure somewhat longer than the duration of the motion in the experimental period. Secondly, the subjects reported little change in their own feelings of well-being. Thirdly, in the tracking task there was no change in performance over a period of about an hour. Were nausea an important factor, performance would decline with time as nausea increased.

This study is clearly preliminary. It involves the dynamic response of only one sort of ship to one sea-state. However, in general, it confirms the

findings of Jex et al. with a very different sort of ship under much rougher conditions. They found very little change in performance of a desk-top calculator task, a consistent 20-40% drop in tracking performance and breakdown in a task requiring unsupported arm movements. The major point of difference between the studies is that Jex et al. dismiss motion under $1m/s^2$ r.m.s. as unlikely to affect performance. It is clear from our study that there are reliable changes in performance well below $1m/s^2$ r.m.s.

There has been remarkable little work to date on the influence of ship motion on manual control skills. It seems necessary to demonstrate two conditions before it is worth investing a major human factors effort in designing the man/control interface on board ship to minimize the effects of motion. Firstly it needs to be shown that some tasks are much more affected by motion than others, otherwise there would be no scope for optimizing the design around tasks which are relatively unaffected by motion. Secondly, it is necessary to show that nausea is not the major determinant of the decrement. If it were, this would indicate a job for a pharmacologist rather than an ergonomist. This paper has demonstrated that both these conditions can be met.

As an example of the importance of these results we might consider the design of a system to allow an observer to identify a point of interest on a radar display on board ship. A recommendation from existing human factors wisdom, which all derives from land based experiments, would suggest, other things being equal, that a light-pen was preferable, a joy-stick controlled cursor next best and a keyboard entry specifying the appropriate matrix point on the display the least efficient. It is clear from the results of our experiments that the results of land based experiments cannot simply be transferred to ships, for a light pen would be the most affected by ship motion and keyboard entry the least. Of course, this does not mean that light pens must not be used in moving environments. But it does mean that a proper programme of research should be mounted to investigate the effects of likely movements to be met under operational conditions on the tasks in question before devices which involve unrestrained limb movement are used in moving environments.

ACKNOWLEDGMENTS

The authors wish to thank Harry Hughes and the staff of the Warren Spring simulator, who did so much to make this work possible. The research was funded by the Office of the Senoir Psychologist (Naval) and coordinated by Hugh Stockbridge. The authors received considerable help in running the experiments and analysing the data from Bob Edwards, Dave Howarth and Bob Olsson. Geoff Allen helped with some of the technical analysis. Finally, we would like to thank the members of the APU staff who volunteered to act as subjects.

APPENDIX

Description of the Warren Spring ship motion simulator

The WSL ship's motion simulator is a counter-balanced gimbal mounted platform system with a heave displacement capability of 3·2m and out-to-out angular motion of 14° in both roll and pitch axes.

The heave motion is derived from a servo-controlled hydraulic motor and reduction gear which drives the payload and counterbalance platforms up and down each side of a central supporting mast structure. The two platforms are coupled by a chain which hangs over the drive sprocket on the main shaft at the top of the mast. Thus the platforms are operated as a balanced system and backlash is minimised. Auxiliary free running cables and pulleys are also provided as a safety measure in the event of mechanical failure in the main chain.

Roll and pitch motions are obtained by servo-controlled hydraulic piston actuators acting against the gimbal mounted platform. These are double acting pistons which move through ± 76mm (3 in) to give the out-to-out displacement of 14°.

The motions of the simulator are controlled (1) by locally generated sine-wave signals from which it can be programmed for either single or combined motions and (2) by external signals which includes recorded ship's motion data or synthesized random data. Operation by sine-wave signals allows individual control of frequency, amplitude and phase relationship for all three motions. To prevent possible damage to the simulator, external signals are connected through a low-pass filter and attenuator to limit signals to within safe operating capability of the simulator.

Collins, A. M., 1973, Decrements in tracking and visual performance during tracking. Human Factors, 15, 379-393.

Drennen, T. G., Curtin, J. G., and Warner, H. D., 1977, Manual control in tracking tasks as a function of control type, task loading and vibration. McDonnel Douglas Astronautics Company--Report No: MDC E1713.

Guignard, J. C., and King, P. F., 1972, Aeromedical aspects of vibration and noise. AGARD-AG-51.

Jex, H. R., O'Hanlon, J. F., and Ewing, C. L., 1976, Simulated rough water operations during long cruises in a 2000-ton surface effect ship, phases 1 and 1A. Systems Technology, Ins., Technical Report, 1057-2.

O'Hanlon, J. F., and McCauley, M. E., 1974, Motion sickness incidence as a function of the frequency and accelaration of vertical sinusoidal motion. Aerospace Medicine, 45, 366-369.

O'Hanlon, J. F., Miller, J. C., and Royal, J. W., 1976, Simulated rough water operations during long cruises in a 2000-ton surface effect ship--Phase 2. Human Factors Research, Inc., Technical Report, 1757-2.

Reason, J. T., and Graybiel, A., 1969, Changes in subjective estimates of well-being during the onset and remission of motion sickness symptomatology in the slow rotation room. NASA-NAMI joint report: NAMI-1083.

THE ROLE OF MANIPULATOR CHARACTERISTICS IN SELECTING THE IDEAL "EFFECTIVE VEHICLE"

By Ronald A. Hess

NASA Ames Research Center

SUMMARY

A structural model of the human pilot has been introduced and discussed in recent Manual Control Conferences. The model has been used to provide a rationale for certain nonlinear pilot control behavior such as stick pulsing and has served as a framework for studying aspects of motor skill development. In light of the theoretical background provided by the model, some past empirical pilot response phenomena are analyzed and shown to be attributable to manipulator or control stick characteristics. In particular, some recent problems associated with pilot/vehicle performance in glideslope tracking in short-takeoff and landing (STOL) aircraft are analyzed. The apparent contribution of the cockpit manipulator (throttle) characteristics to these problems are outlined and a solution proposed and evaluated in both simulation and flight test.

INTRODUCTION

In order to actively control some physical system such as an aircraft or automobile, the human operator must utilize a manipulator such as a control stick or steering wheel. The characteristics of the manipulator can have a profound effect upon the performance of the man-machine system. Although studies such as those by Herzog[1] and Merhav and Ya'Acov[2] have capitalized upon this interface to improve tracking performance in certain compensatory tasks, the specific inclusion of manipulator characteristics has not been a primary concern of the analyst. In pilot modeling, for example, the human has generally been treated as a servomechanism with zero output impedance. If the dynamics of the manipulator are significant in the frequency range of interest for manual control (typically $0.1 < \omega < 10$ rad/sec), they are usually lumped into those of the controlled element. Proprioceptive feedback has been postulated to fulfill a relatively minor role in determining overall pilot input-output characteristics although its contribution to the operation of the particular neuromuscular system operating the manipulator has been recognized as extremely important.[3]

Recently, Hess[4-7] has introduced and discussed what can be called a structural model of the human pilot in which proprioceptive feedback plays an important role in determining pilot equalization. In the next section,

this model will be discussed briefly and the implication of its structure as regards manipulator characteristics will be treated. With the model serving as a theoretical framework, some specific empirical examples of manipulator effects will then be discussed.

THE STRUCTURAL MODEL

The structural model of the human pilot proposed by Hess has been discussed at some length in the literature[4-7] and hence will only be outlined here. Figure 1 is a block diagram of this model for compensatory tracking behavior. The model of Fig. 1 has been divided into "central nervous system" and "neuromuscular system" components, a division intended to emphasize the nature of the signal-processing activity involved. System error $e(t)$ is presented to the pilot via a display with dynamics Y_{d_e}. The rate of change of the displayed error is assumed to be derived from $e_d(t)$. The process of deriving error-rate is assumed to entail a computational time delay of τ_1 seconds. Constant gains K_e and $K_{\dot{e}}$ multiply the signals $e_d(t)$ and $e_d(t-\tau_1)$, respectively. The switch allows either of these two signals to be used as driving signals to the remainder of the model. A discussion regarding the utility of error-rate control is provided in Ref. 6. The action of the switch is parameterized by the variable P_1, which represents the probability that the switch will be in position 1 (error-rate control) at any instant of time. A central time delay of τ_0 seconds is included to account for the effects of latencies in the visual process sensing $e_d(t)$, motor nerve conduction times, etc. The resulting signal $u_c(t)$ provides a command to a closed-loop system, which consists of a model of the open-loop neuromuscular dynamics of the particular limb driving the manipulator, Y_{p_n}, and elements Y_f and Y_m, which emulate, at least approximately, the combined effects of the muscle spindles and the dynamics associated with higher level signal processing. A colored noise $n_u(t)$ is injected at the pilots's output as remnant.

As pointed out in Ref. 6, the signal $u_m(t)$ is really proportional to the time rate of change of vehicle output due to control activity, and as such, is a form of rate feedback. The first three rows of Table 1 show model parameters selected to give the describing function matches shown in Figs. 2-4. Table 2 shows the variation in pilot dynamics (in simplified form) with increases in the order of the controlled element dynamics. The third column shows the simplified form of the proprioceptive feedback implied by the combination of $Y_f Y_m$ in Fig. 1. For example, for $Y_c = K/s$, $k=1$, and from Fig. 1,

$$Y_f Y_m = K_1 s/(s + 1/T_1) \qquad (1)$$

For values of T_1 found appropriate for K/s dynamics ($T_1 \doteq 5$ secs from Table 1), $Y_f Y_m$ looks very much like a pure gain in the important region of open-loop crossover. Thus, row 2, column 3 of Table 2 shows the required proprioceptive feedback for controlling K/s dynamics to be applied force or displacement $u_\delta(t)$. This force or displacement is defined relative to a set-point or regulation point, e.g., the equilibrium position of a spring-

restrained control stick. K/s dynamics have long been associated with the most desirable "effective vehicle" characteristics for single-axis systems under manual control.[8] In terms of the classical servo-model of the human pilot (likened to column 2 of Table 2), a "pure-gain" pilot results, i.e., no pilot equalization is required. In terms of the structural model, which inherently contains two feedback loops, only feedback of proprioceptively sensed force or displacement is needed. The same cannot be said for $Y_c=K/s^2$. Here, k=2 and from Fig. 1

$$Y_f Y_m = K_1 s/(s + 1/T_1)(s + 1/T_2) \qquad (2)$$

For values of T_1 and T_2 found appropriate for K/s^2 dynamics ($T_1=T_2\doteq 2.5$ secs from Table 1) $Y_f Y_m$ looks very much like an integrator in the important region of crossover. Thus, row 3, column 3 of Table 2 shows the required proprioceptive feedback for controlling K/s^2 dynamics to be the integral of applied force or displacement from some set-point or regulation-point. A rationale for human operator pulsive control behavior was offered in Ref. 5 based upon the hypothesis that the human attempts to reduce the computational burden of time integration of $u_\delta(t)$ in higher levels of the central nervous system.

Lastly, consider $Y_c=K$, k=0 and, from Fig. 1,

$$Y_f Y_m = K_1(s + 1/T_2)s/(s + 1/T_1) \qquad (3)$$

With $T_1= T_2$, $Y_f Y_m$ takes the form of a differentiator. Thus, row 1, column 3 of Table 2 shows the required proprioceptive feedback for controlling K dynamics to be the time derivative of $u_\delta(t)$. Again, assuming the validity of the model of Fig. 1, this differentiation might also be accompanied by considerable activity in the higher levels of the central nervous system. Two things may mollify this situation, however. First, as Fig. 1 and Table 2 indicate, the pilot dynamics u_δ/e_d for this controlled element are a first order lag. This inherent filtering action of the error signal makes the pilot output $u_\delta(t)$ rather smooth and low frequency in nature. This is exemplified in Figs. 5 and 6 taken from Ref. 5 where the structural model was digitally simulated as part of a single-axis tracking task. Fig. 5 shows segments of $e_d(t)$ and $u_\delta(t)$ for $Y_c=K$ whereas Fig. 6 shows the same variables for $Y_c=K/s$. Notice the lower frequency content of $u_\delta(t)$ in Fig. 5 as opposed to that in Fig. 6. Second, the muscle spindles and Golgi tendon organs themselves, can provide direct rate information. This means that differentiation as an <u>operation</u> in the higher levels of the central nervous system may be obviated. Of special importance is the fact that the required proprioceptive feedback (or calculation) of du_δ/dt does <u>not</u> require information regarding a set-point or regulation point as was the case in the previous two controlled elements ($Y_c=K/s,K$). This will have important repercussions in the section which follows.

FLIGHT PATH CONTROL OF STOL VEHICLES

Reference 10 summarizes some interesting work, part of which involved the landing approach performance of a simulated powered-lift short takeoff and

landing (STOL) aircraft. Various vehicle dynamics were evaluated in piloted simulation. In the landing approaches, vertical flight path control was accomplished almost exclusively by throttle. In addition, very little column activity was needed for attitude/airspeed control. This was not accidental as an attitude-hold stability augmentation system (SAS) was designed and utilized for the express purpose of minimizing pilot activity with the longitudinal control column.

Figure 7 shows the dynamics of one of the configurations analyzed. The pertinent transfer function is

$$(d/\delta_T)' = (N^{\dot{d}}_{\delta_T})'/s\Delta' \qquad (4)$$

Here, d represents longitudinal vehicle motion perpendicular to the glideslope, and δ_T is throttle movement. The (') notation is meant to emphasize the fact that an inner attitude-loop is being closed by the SAS. No pilot inner-loop attitude closure is assumed, an assumption found to be valid from simulation. Table 3 lists the pertinent vehicle dynamics.

Figure 8 shows pilot/vehicle transfer function for the configuration of Fig. 7 measured at six frequencies around crossover. This "open loop" transfer function is of interest for four reasons: First, the crossover frequency appears to be somewhere between 0.3 and 0.4 rad/sec, a very low value for manual control experiments. Second, the pilot is not particularly successful in forcing the open-loop pilot/vehicle characteristics into a K/s-like form in the region of crossover. Third, the low frequency phase data exhibits none of the "phase droop"[6] normally associated with such pilot/vehicle open loop transfer function measurements. Finally, fitting this data with a simple lead-lag model would require an effective time-delay of 0.8 secs, quite a large value for manual control experiments.

We will now show that these four characteristics can be produced by the structural model of Fig. 1. Figure 9 shows the model-generated pilot/vehicle transfer function. The model parameters are listed in the fourth row of Table 1. The model fit was obtained by assuming that the pilot was controlling <u>rate</u> alone. This assumption was necessary to achieve an acceptable fit to the data. Actually, of course, the error-rate loop would serve as an inner-loop to an outer, error-loop closure. However, the fact that a reasonable fit to the data could be obtained by considering just error-rate control, alone, suggests that this control dominates. This is corroborated by experimental results from Ref. 10 where it was stated "All of the pilot's indicated that the technique for glideslope tracking was primarily to control glideslope deviation <u>rate</u> ($\dot{d}$)."

Using the structural model, we can now provide a rationale for this activity. The transfer function $(\dot{d}/\delta_T)'$ will exhibit pure-gain like characteristics for $\omega \leq 0.3$ rad/sec. Such characteristics have been hypothesized here to be ideal for manipulators which do not provide set-point information, such as the engine throttles used in Ref. 10. Normally, exclusive rate control would carry a workload burden emposed by the necessity of deriving rate information from displacement information. However, in the simulation of Ref. 10, rate information was available <u>directly</u>

from the Instantaneous Vertical Speed Indicator (IVSI) in the cockpit. Again, quoting from Ref. 10, the actual piloting technique was

"a. Keep $\dot{d}$ at a very low level by controlling IVSI with power, e.g., find a target IVSI that keeps the glideslope bug stationary on the display (nominally 800 ft/min)

b. If glideslope error (d) is diverging, try to first zero $\dot{d}$, than adjust power so d is slowly converging (i.e., pick a new target sink rate on the IVSI).

c. If the glideslope error is less than one dot, make very small power adjustments (if any)."

Since rate information was available directly, the delay normally associated with rate derivation, τ_1 , was set to zero. To account for scanning delays, τ_0 was increased from the nominal 0.14 secs to 0.2 secs. Note that the model captures the salient features of the data including the four "anomolies" mentioned previously. In particular, note that no large time delays have to be hypothesized to match the phase lag data. It also appears that all of these anomolies have their origin in the characteristics of the manipulator and in the availability of explicit rate information.

Next, let us consider the results of an investigation reported in Ref. 11. In this study a flight test program was carried out to assess the feasibility of piloted instrument approaches along pre-defined, steep, curved and decelerating approach profiles in powered-lift aircraft operating on the backside of the power curve. Separate stability augmentation systems for attitude and speed were provided, as well as a supporting flight director and special electronic cockpit displays. Of particular interest was a problem encountered in glideslope tracking using a throttle flight director which produced K/s-like effective-vehicle characteristics in the frequency range $0.1 \leq \omega \leq 1.0$ rad/sec. Figure 10 shows the effective-vehicle characteristics (director+aircraft). Although the K/s dynamics do not extend beyond 1.0 rad/sec, no additional and deleterious phase lags accrue in this region. Figure 11 shows the oscillatory glideslope tracking characteristics revealed in flight tests for this configuration. The question now arises as to how the pilot would control this effective vehicle. Two obvious approaches are: 1) use rate control as in the previous example, 2) use displacement control. In the first case, the pilot's internal model of the effective vehicle in the frequency range $\omega \leq 1.0$ rad/sec would be a pure gain (k=0). According to the structural model, this would allow the manipulator to be suited to the dynamics. However, unlike the simulation just studied, rate information in the form of $\dot{\delta}_{TFD}$ (rate of change of throttle flight-director signal) is not explicitely available and would have to be derived by the pilot. Probable pilot/vehicle dynamics for this case are shown in Fig. 12. The structural model parameters are shown in the fifth row of Table 1. With the exception of τ_0 and τ_1 , they are identical to the parameters which yielded the match of Fig. 9, (fourth row of Table 1).

The value of τ_0 was increased from 0.2 to 0.5 secs to account for the fact that considerably more scanning probably occurred in the study reported

in Ref. 11 (a flight test) than in the one reported in Ref. 10 (a simulation). The value of τ_1 was increased from 0 to 0.2 secs to account for the fact that rate information had to be derived. As Fig. 12 indicates, stability margins are more than adequate. However, the necessity of continuously deriving rate information from the displayed flight-director signal δ_{TFD} may lead to high workload and inadequate time to control the remaining two directors and scan the status displays in the cockpit. In addition, rate control alone may not yield performance which the pilot deems acceptable.

Consider, on the other hand, the situation where the pilot controls displacement. Here, the pilot's internal model of the effective vehicle in the frequency range $\omega \leq 1.0$ rad/sec would be K/s (k=1). According to the structural model, the manipulator characteristics are not well suited to those of the effective-vehicle. It has been hypothesized here that the necessary proprioceptive information of applied force or displacement from a set-point (see Table 2) would not be available. Figure 13 shows the probable pilot/vehicle dynamics for this case under the preceding assumption. The model parameters are given in the sixth row of Table 1. They are closely related to those of the second row which utilized K/s dynamics. The most significant differences are the values of T_1 and τ_0. The rationale behind the revised value of τ_0 has just been given. For the sake of simplicity, no switching is assumed to occur, i.e., $P_1=0$. The decreased value of T_1 is intended to account for the assumption that little low-frequency proprioceptive feedback is available from the overhead throttles. This reduction means that the "break frequency" of the washout element Y_f of Fig. 1 is moved to higher frequencies, thus reducing the amount of low-frequency information available.

The effect of this change is rather dramatic as can be seen by comparing the phase angle plots of Figs. 12 and 13. Note the much larger phase lags apparent in Fig. 13. This lag increment is obviously not due to any changes in the delay τ_0, however. Rather, the closure of the two inner-loops of the structural model cause a real root to migrate to a position $s \doteq -0.4$. This is demonstrated in the two root locus diagrams for these closures shown in Figs. 14 and 15. Note from Fig. 13 that a closed loop instability is possible at $\omega=0.8$ rad sec (0.13 cycles/sec). This is seen to compare quite favorably with the frequency of the path rate oscillations evident in Fig. 11. These oscillations constitute the glideslope tracking problem alluded to briefly at the beginning of our discussion of the experiments of Ref. 11. Although the oscillations of Fig. 11 represent a worst case example, they typified the glideslope tracking characteristics of the pilot/vehicle system. In Fig. 11, 5 cycles occur in approximately 38 secs (0.13 cycles/sec). The model results should not be interpreted as a "prediction" but rather as a rationale for the existence of a low frequency oscillation in the glideslope tracking for this effective-vehicle/manipulator combination. Note that in order to produce an unstable frequency at 0.8 rad/sec using a crossover model of the pilot,[8] an effective time delay on the order of 1.5 secs would have to be hypothesized!

From what has been discussed thus far, a solution to the flight-path oscillation problem would appear to lie in changing the characteristics of

either the throttle or the effective vehicle. The latter course was chosen in Ref. 11 and the dynamics of the effective vehicle (director+aircraft) were changed from K/s to K. This was accomplished by feeding back washed-out throttle position to the director. The reader is referred to Ref. 11 for details. Figure 16 shows the modified effective vehicle dynamics. The roll-off at frequencies beyond 5 rad/sec is due to a first-order filter being implemented to smooth the director signal at high frequencies. Figure 17 shows the resulting flight test results with the modified director (note the change in scales in the ordinates between Figs. 11 and 17). Performance is improved rather dramatically. Quoting from Ref. 11:

> "During the limited flight evaluation of these alternative throttle flight director control laws, and during the course of gathering the simulator data, pilot commentary indicated very little tendency toward oscillatory glidepath tracking characteristics. The pilots were not aware of providing any compensation while tracking the throttle-director bar, and were able to easily null the flight-director command bar without overshoot, using what were frequently step-like throttle inputs as can be seen in figure 63 (our Fig. 17). Once a correction was made, attention could temporarily be diverted to other display-scanning tasks without large errors developing in the throttle-director bar"

CLOSING REMARKS

The research just discussed was intended to point out that "ideal" effective-vehicle dynamics for manual control systems can be dependent upon manipulator characteristics. A structural model of the human pilot described in Ref. 6 which incorporated explicit proprioceptive feedback was used as a framework for interpreting some simulation and flight test results. The model was able to match measured pilot transfer function data exhibiting "anomolous" characteristics and was able to provide a rationale for oscillatory pilot/vehicle behavior. Finally, the model suggested a successful approach for improving the glideslope tracking performance of the aircraft exhibiting the oscillatory behavior.

REFERENCES

1. Herzog, J. H., "Proprioceptive Cues and Their Influence on Operator Performance in Manual Control," NASA CR-1248, 1969.

2. Merhav, I. J., and Ya'Acov, O. B., "Control Augmentation and Work-Load Reduction by Kinesthetic Information from the Manipulator," IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-6, No. 12, Dec. 1976, pp. 825-835.

3. Magdaleno, R. E., and McRuer, D. T., "Experimental Validation and Analytical Elaboration for Models of the Pilot's Neuromuscular Subsystem in Tracking Tasks," NASA CR-1757, 1971.

4. Hess, R. A., "A Dual-Loop Model of the Human Controller, "Journal of Guidance and Control, Vol. 1, July-Aug., 1978, pp. 254-260.

5. Hess, R. A., "A Rationale for Human Operator Pulsive Control Behavior," Journal of Guidance and Control, Vol. 2, May-June 1979, pp. 221-227.

6. Hess, R. A., "A Structural Model of the Adaptive Human Pilot," Journal of Guidance and Control, Vol. 3, No. 5, Sept.-Oct. 1980, pp. 416-423.

7. Hess, R. A., "Pursuit Tracking and Higher Levels of Skill Development in the Human Pilot," IEEE Transactions on Systems, Man, and Cybernetics, April 1981.

8. McRuer, D., "Human Dynamics in Man-Machine Systems," Automatica, Vol. 16, No. 3, May 1980, pp. 237-253.

9. Granit, R., The Basis of Motor Control, Academic Press, London and New York, 1970.

10. Hoh, R. H., Craig, S. J., and Ashkenas, I. L., "Identification of Minimum Acceptable Characteristics for Manual STOL Flight Path Control," Federal Aviation Administration, FAA-RD-75-123, June 1976.

11. Hindson, W. A., Hardy, G. H., and Innis, R. C., "Flight-Test Evaluation of STOL Control and Flight-Director Concepts in a Powered-Lift Aircraft Flying Curved Decelerating Approaches," NASA Technical Publication 1641, 1981.

Table 1. Structural Model Parameter Values

Controlled-element dynamics	Model parameters											
	k	K_e	K_e'	K_2	P_1	T_1	T_2	K_1	τ_0	τ_1	ζ_n	ω_n
K	0	11.1	2.13	2.0	0.05	5.0	5.0	1.0	0.14	0.2	0.707	10.0
K/s	1	22.2	3.42	2.0	0.05	5.0	-	1.0	0.14	0.2	0.707	10.0
K/s^2	2	26.2	10.50	10.0	0.20	2.5	2.5	1.0	0.14	0.2	0.707	10.0
$(d/\delta_T)'$	0*	0.0	87.50	50.0	1.00	2.0	0.5	1.0	0.20	0.0	0.707	10.0
δ_{TFD}/δ_T	0*	0.0	87.50	50.0	1.00	2.0	0.5	1.0	0.50	0.2	0.707	10.0
δ_{TFD}/δ_T	1	0.77	0.0	10.0	0.00	0.2	-	1.0	0.50	-	0.707	10.0

*Internal model appropriate for rate tracking

Table 2. The Adaptive Pilot

Controlled Element Y_c	k	Simplified pilot dynamics Y_p	Required proprioceptive feedback
K	0	$K_p e^{-\tau_e s}/(T_I s + 1)$	$du_\delta(t)/dt$
K/s	1	$K_p e^{-\tau_e s}$	$u_\delta(t)$
K/s^2	2	$K_p(T_L s + 1)e^{-\tau_e s}$	$\int u_\delta(t)dt$

Table 3. STOL Configuration AP1 (Ref. 10)

V_0 (kts)	75.0
γ_0 (deg)	-6.0
θ_0 (deg)	1.87
δ_{T_0} (%)	30.6
Δ'	$(s + .25)(s + .36)(s + .67)(s + 8.32)(s^2 + 2(.5)(.33)s + .33^2)$
$(N^{\dot{d}}_{\delta_T})'$	$0.167(s + .475)(s + 1.75)(s + 8.3)(s^2 + 2(.99)(.163)s + .163^2)$

Prime notation indicates inner attitude-loop closed by SAS

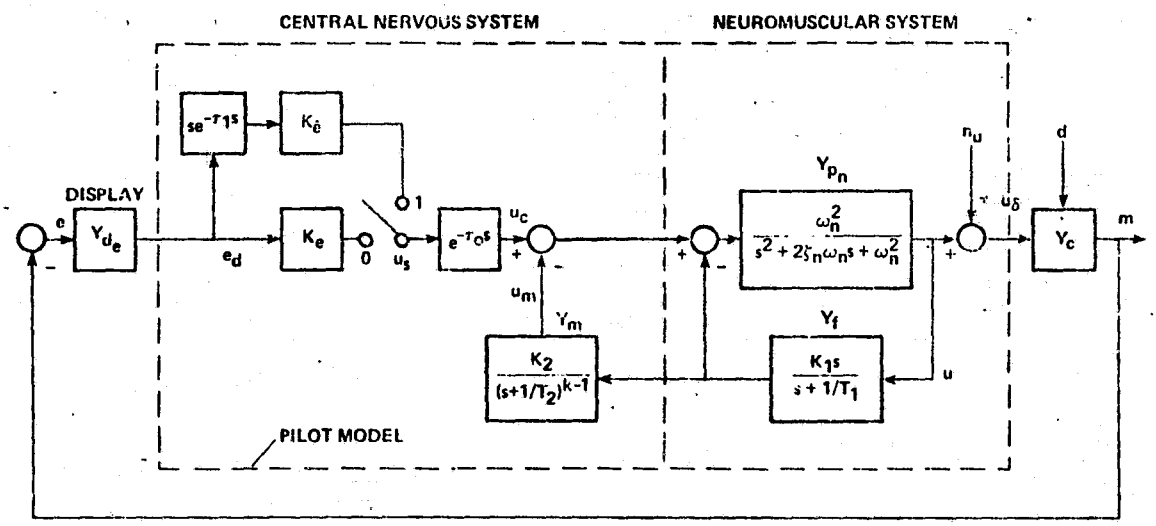


Figure 1. The structural model of the human pilot

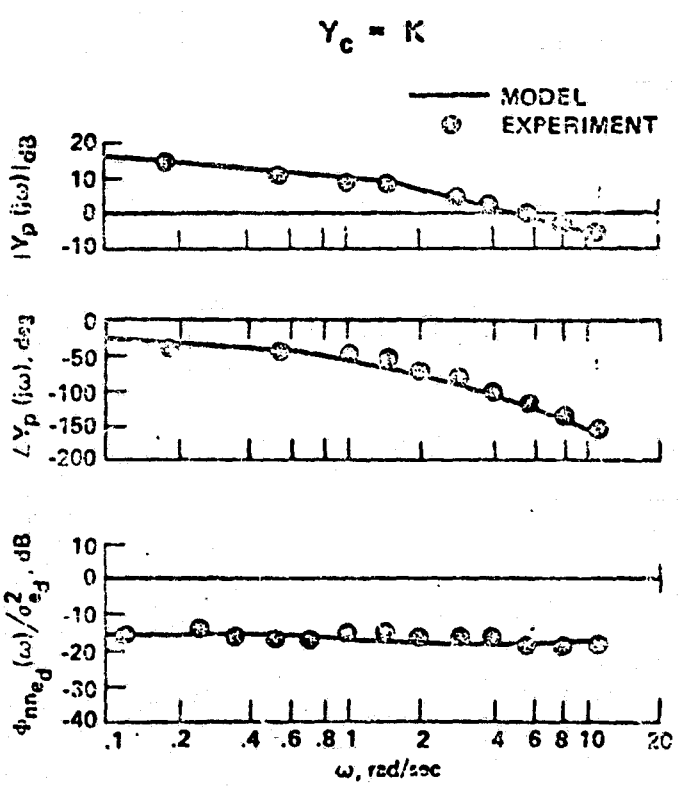


Figure 2. Describing function and remnant comparison, K controlled-element dynamics

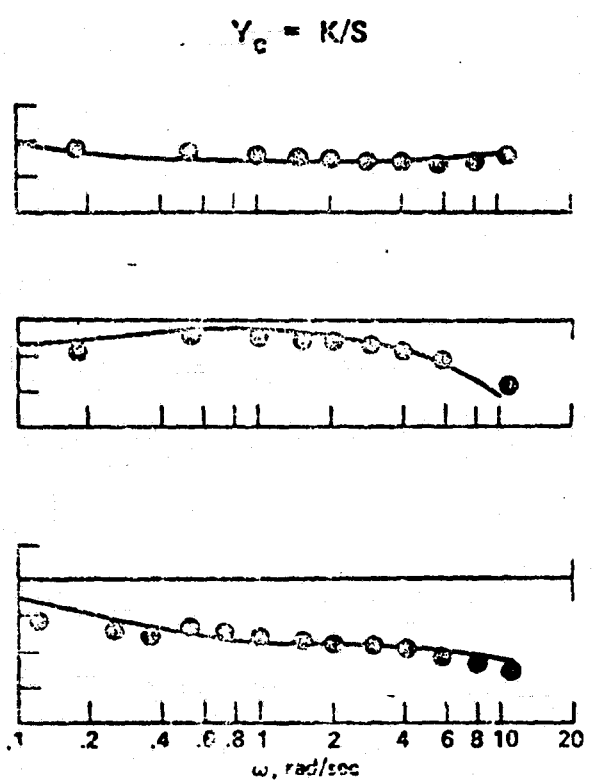


Figure 3. Describing function and remnant comparison, K/s controlled element dynamics

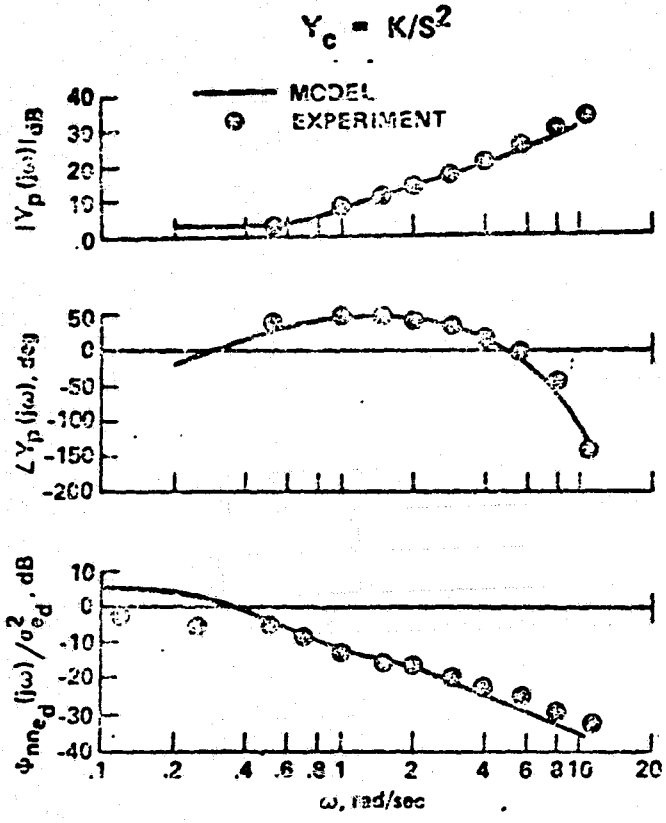


Figure 4. Describing function and remnant comparison, K/s^2 contolled-element dynamics

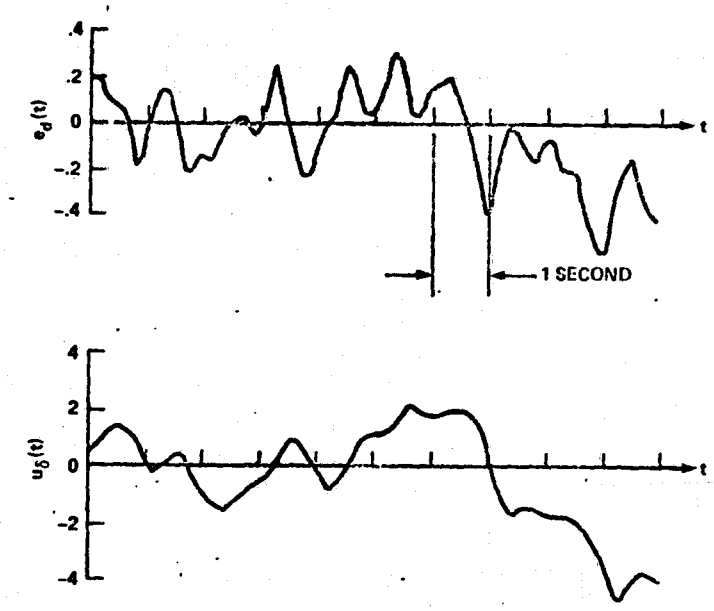


Figure 5. Error and control from structural model, K controlled-element dynamics

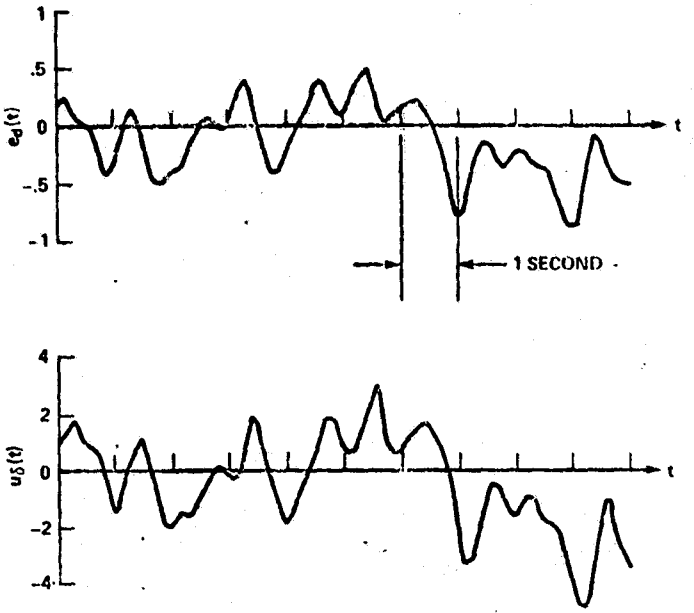


Figure 6. Error and control from structural model, K/s controlled-element dynamics

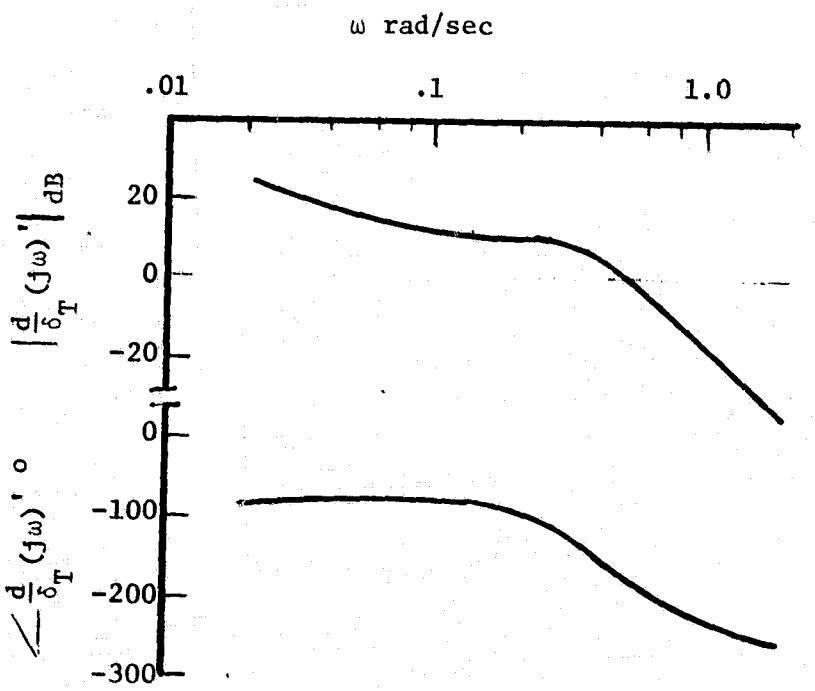


Figure 7. Vertical flight path to throttle characteristics for config. AP1 from Ref. 10

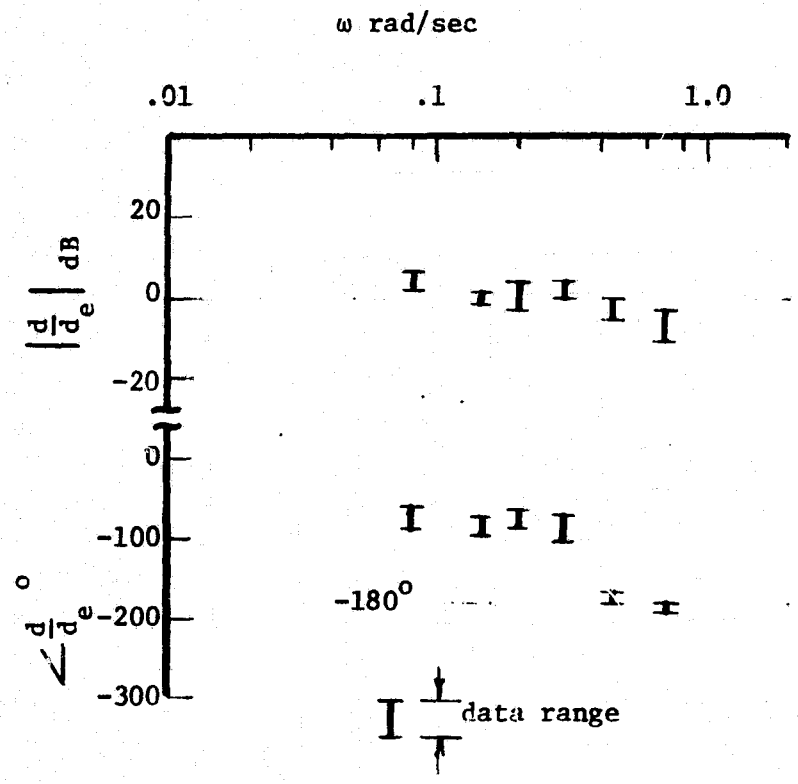


Figure 8. Measured pilot/vehicle transfer function for config. AP1 from Ref. 10.

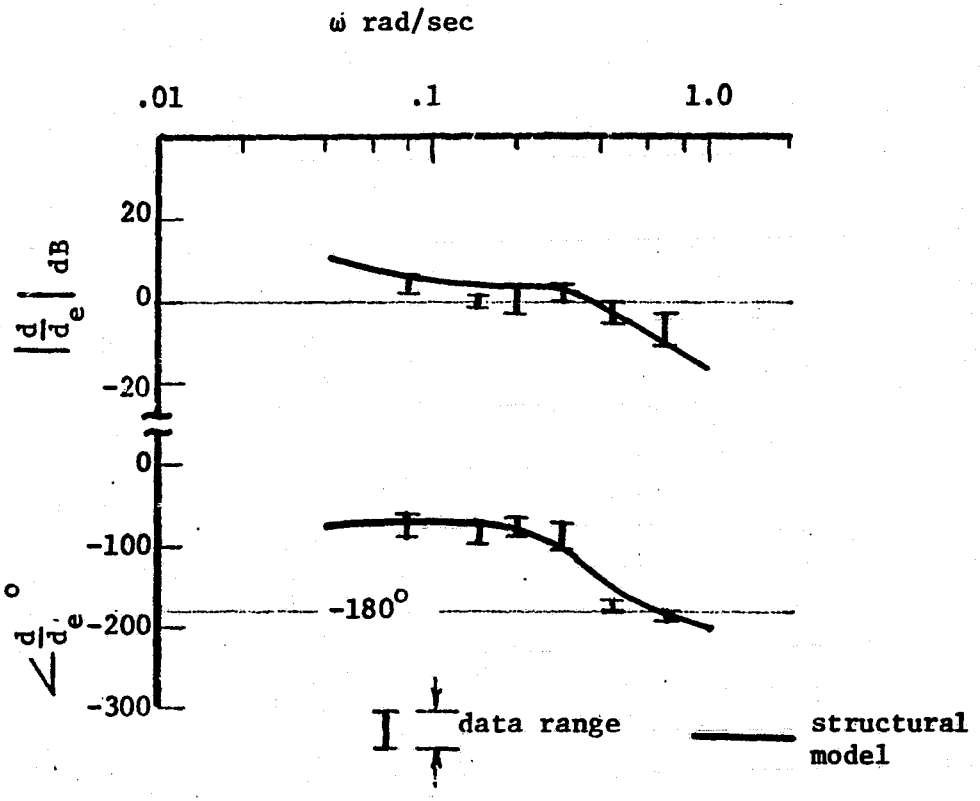


Figure 9. Comparison of experimental and model-generated transfer functions for config. AP1 from Ref. 10

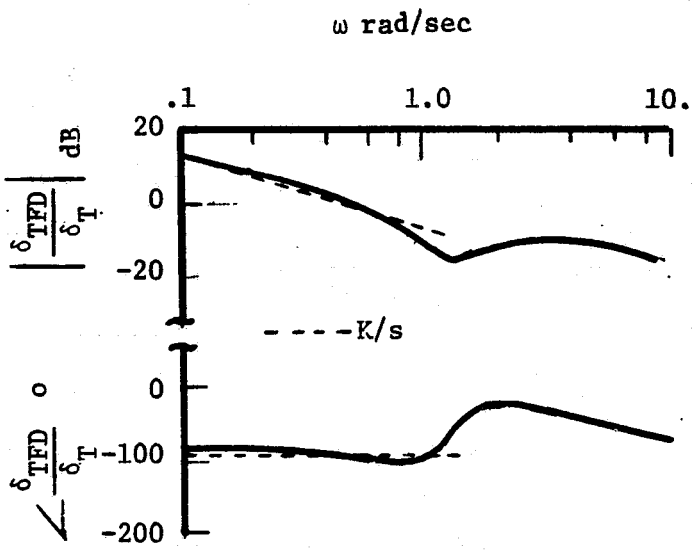


Figure 10. Effective vehicle (director+aircraft) for throttle control from Ref. 11

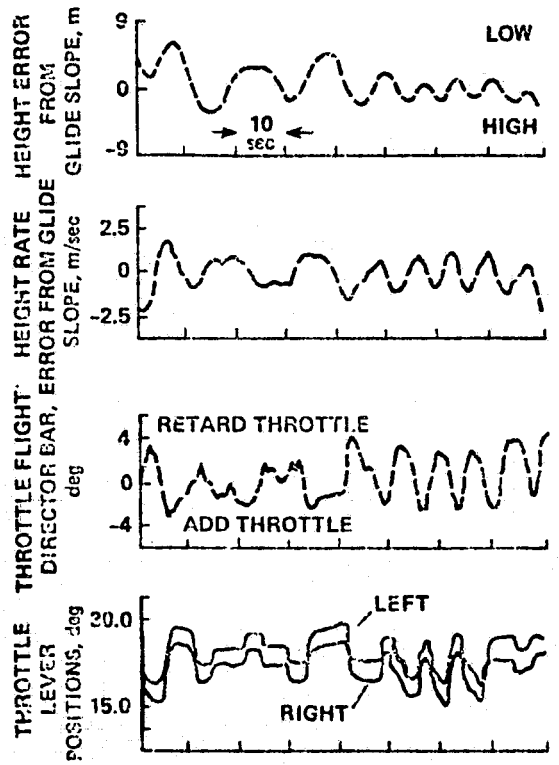


Figure 11. Glideslope tracking characteristics of aircraft from Ref. 11

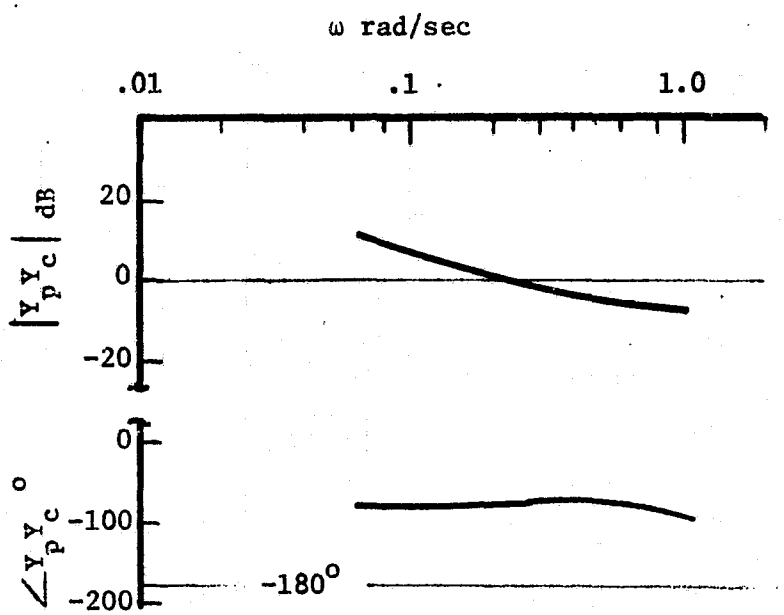


Figure 12. Model-generated transfer function for rate control of effective vehicle of fig. 10

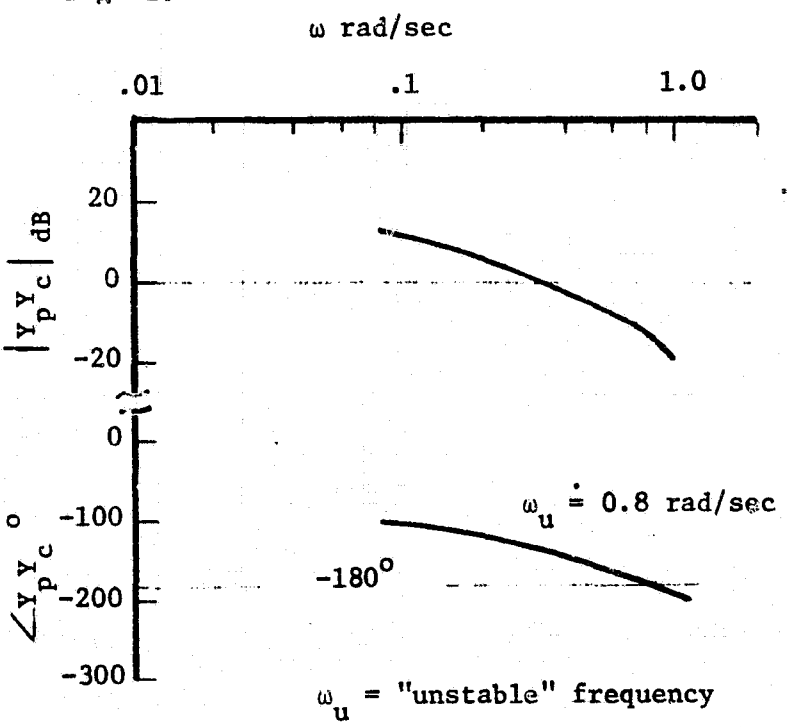


Figure 13. Model-generated transfer function for displacement control of effective vehicle of fig. 10

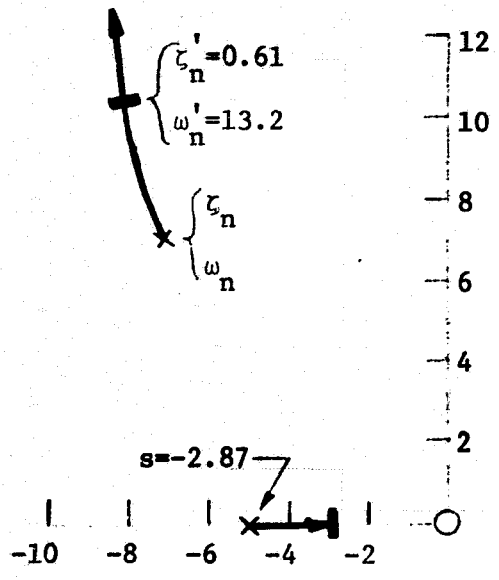


Figure 14. Root locus diagram for structural model inner-loop closure, model characteristics shown in fig. 13

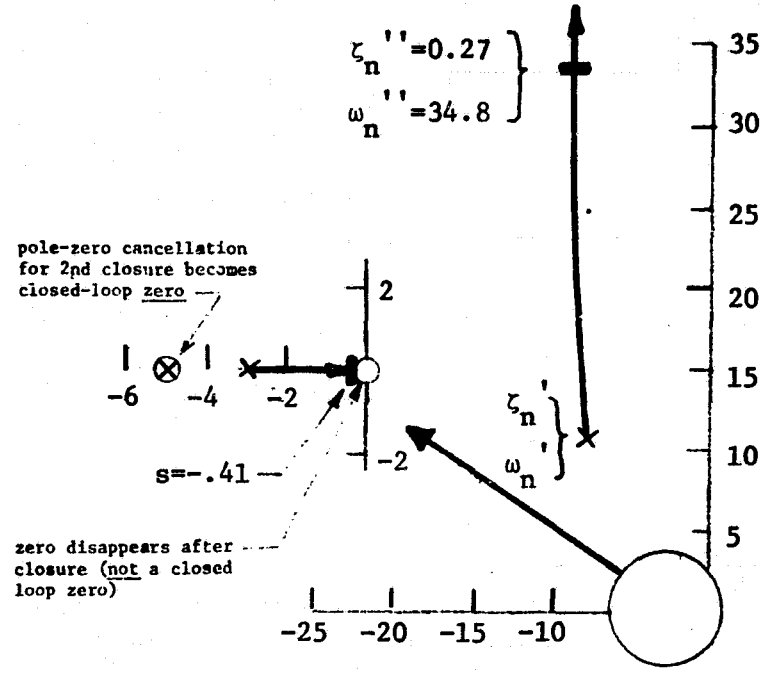


Figure 15. Root locus diagram for structural model outer-loop closure; model characteristics shown in fig. 13

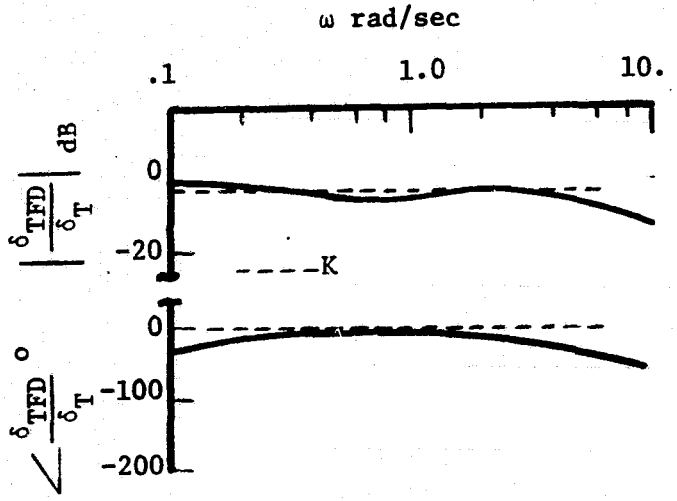


Figure 16. Modified effective vehicle (director+aircraft) for throttle control from Ref. 11

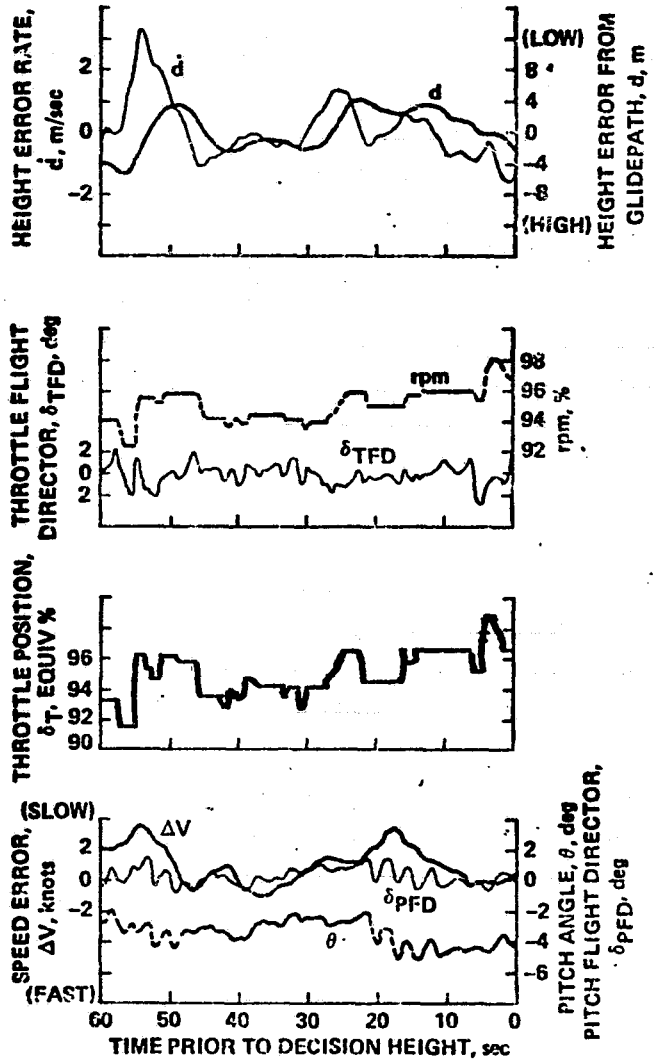


Figure 17. Glideslope tracking characteristics of aircraft with modified flight director from Ref. 11

COMPUTER-AIDED MANUAL TRACKING

By Yung-Koh Yin and Russell F. Berg

General Dynamics, Pomona Division

SUMMARY

A scheme has been developed to assist the human operator by augmenting an optic sight manual tracking loop with target rate estimates from a computer control algorithm which can either be a Kalman Filter or an α, β, γ Filter. The idea is for the computer to provide rate tracking while the human operator is responsible for nullifying the tracking error. A simple schematic is shown to illustrate the implementation of this concept.

A hybrid real-time man-in-loop simulation was used to compare the tracking performance of the same flight trajectory with or without this form of computer-aided track. Preliminary results show the advantage of computer-aided track against high speed aircraft at close range. However, good tracking before target state estimator maturity becomes more critical for aided track than without. Results are presented for a constant velocity flight trajectory.

I. INTRODUCTION

It is a widely accepted fact that the human, while acting as a control element, not only exhibits a variety of nonlinear and time-varying behavior but also posesses unique adaptive and learning capacity. These characteristics make the human an indispensible part of many control and tracking missions. Due to the inherent limitation of a slow reaction time and a "noisy" output of the human tracker, tracking performance deteriorates rapidly when the target trajectory results in high angular rates and accelerations, for example, in the cross-over region near the point of closest approach.

The idea of rate-aided track for an unmanned closed-loop tracking system has long been discussed in detail for example by Fitts (1). However, the application of this concept to the manual tracking system has not been so successful. This paper will present the concept, implementation and simulation of a rate-aided scheme to a manual tracking task.

Section II will show the track loop and the target state estimator (TSE) while section III will show how the aiding rates are generated by the computer, based on the information available in the TSE, as well as how these aided terms are introduced into the track loop controller. Section IV is experiment setup, Section V will present the simulation results for a constant velocity trajectory with and without the rate-aided feedback. Finally, Section VI presents our conclusions and recommendations.

II. OPTIC TRACK LOOP AND TARGET STATE ESTIMATOR

A typical single axis man-in-loop optic track system is depicted in the following block diagram, Figure 1. Whether the track is computer-aided or not depends on the on-off computer-aided switch position.

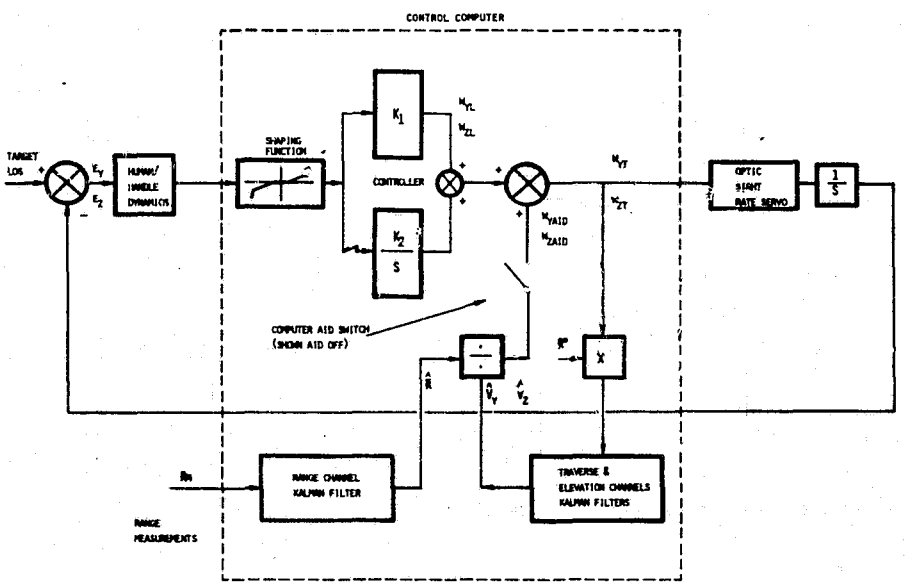


Figure 1. Optic Track Loop and Target State Estimation

Upon seeing an angular tracking error, which is the difference between the target line-of-sight and the optic sight center line, the human tracker outputs a rate command through a thumb transducer to the sight controller. In the present (baseline) case, the controller is a proportional plus integral element in the computer's software and in turn commands the sight servo to follow the target motion. Without any rate estimates to aid the human tracker, the target state estimator is outside the track loop and serves the function of generating target's, velocity and acceleration. It is this information that will be used to generate rate-aided feedback terms.

III. COMPUTER-AIDED RATE FEEDBACK

Based on the range measurement and the two angular rates of the optic sight, the target state estimator estimates the target velocity and acceleration at the end of the digital cycle in the instantaneous sight frame, under the assumption that the tracking errors are small. The outputs of the filtered quantities are range (R*) end point velocity estimates (V_X*, V_Y*, V_Z*) and acceleration (A_X*, A_Y*, A_Z*). In order to generate the proper rates for augmentation to the track loop, the end point velocity estimates are first extrapolated to obtain the future target velocity one computer cycle time ahead and then rotated into the new instantaneous sight frame at the next cycle to obtain $\hat{V}_Y$ and $\hat{V}_Z$. Two filtered angular rates ($\hat{V}_Y/\hat{R}$) and ($\hat{V}_Z/\hat{R}$) are applied to the controller. This is the rate-aided term generation sequence in steady state. Since the filter takes a finite time to settle, including an aid term obtained from an unsettled filter can only degrade tracking performance. Even when the filter is fully matured, the abrupt adding of an aid term is too large a disturbance for the human operator to handle. Therefore, this computer-aided term is ramped into the track loop in about 2-3 seconds while the integral element in the controller is ramped out. We have learned from simulation experiments that good tracking and earlier filter maturity are critical to the tracking performance for aided track. A separate switch, therefore is provided for the human tracker to decide when to use this computer-aided tracking mechanism.

IV. HYBRID COMPUTER EXPERIMENT SET-UP

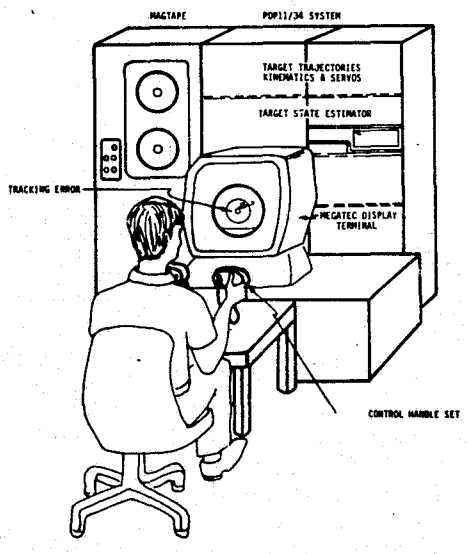


Figure 2. Real-Time Man-in-Loop Simulation

C-6

The simulation facility as shown in Figure 2 consists of a PDP11/34 digital computer, a Megatek graphics display unit, a control handle which interfaces with the host computer via a DEC AR11 A/D converter and various supporting devices. The numerical computation, including system kinematics, servo response, coordinates transformation and tracking error generation, is done by the PDP11/34. The whole simulation is run in real time with a cycle time of 36 msec. The outputs, which are essential to the performance study, are stored in the main memory as they are computed and transferred later onto the disk. The predesigned target trajectories are stored on the disk and loaded into memory before the real time run starts. This is also true for all the initialization of the simulation as well as target display graphics. Whenever the tracking error is greater than the half field view angle, the run is terminated to simulate loss of lock.

V. SIMULATION RESULTS

The target model used in the simulated example is an incoming constant velocity aircraft flying at about 250 m/sec with the closest approach about 300 meters. The commanded rate outputs for human operator for both unaided and aided cases are shown in Figure 3. Similarly, the tracking errors are shown in Figure 4. The tracking is done by the same operator. It should be noted that the system loses lock near the cross-over for the unaided case.

VI. CONCLUSION AND RECOMMENDATION

1. The computer-aided track is not needed for low speed target but does help the operator to track a high speed target in the cross-over region.

2. Early tracking accuracy is essential for the target state estimator to have enough data of good quality to generate aided rates.

3. Research and development in this area will result in improved manual tracking performance.

VII. REFERENCE

1. Fitts, John M.: Aided Tracking as applied to High Accuracy Pointing Systems. IEEE Transactions on Aerospace and Electronic Systems. Vol. 9, No. 3, May 1973.

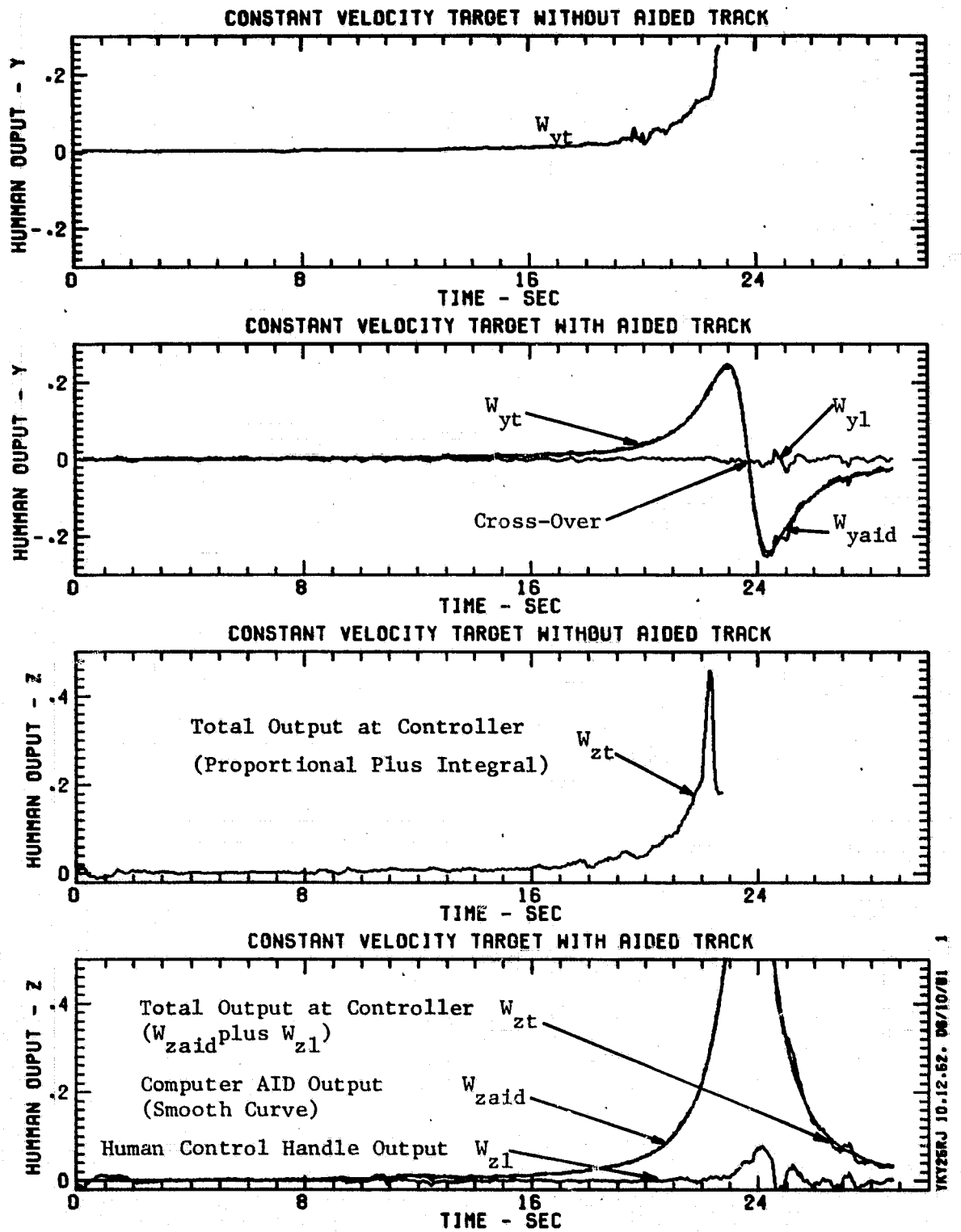


Figure 3. Human Output

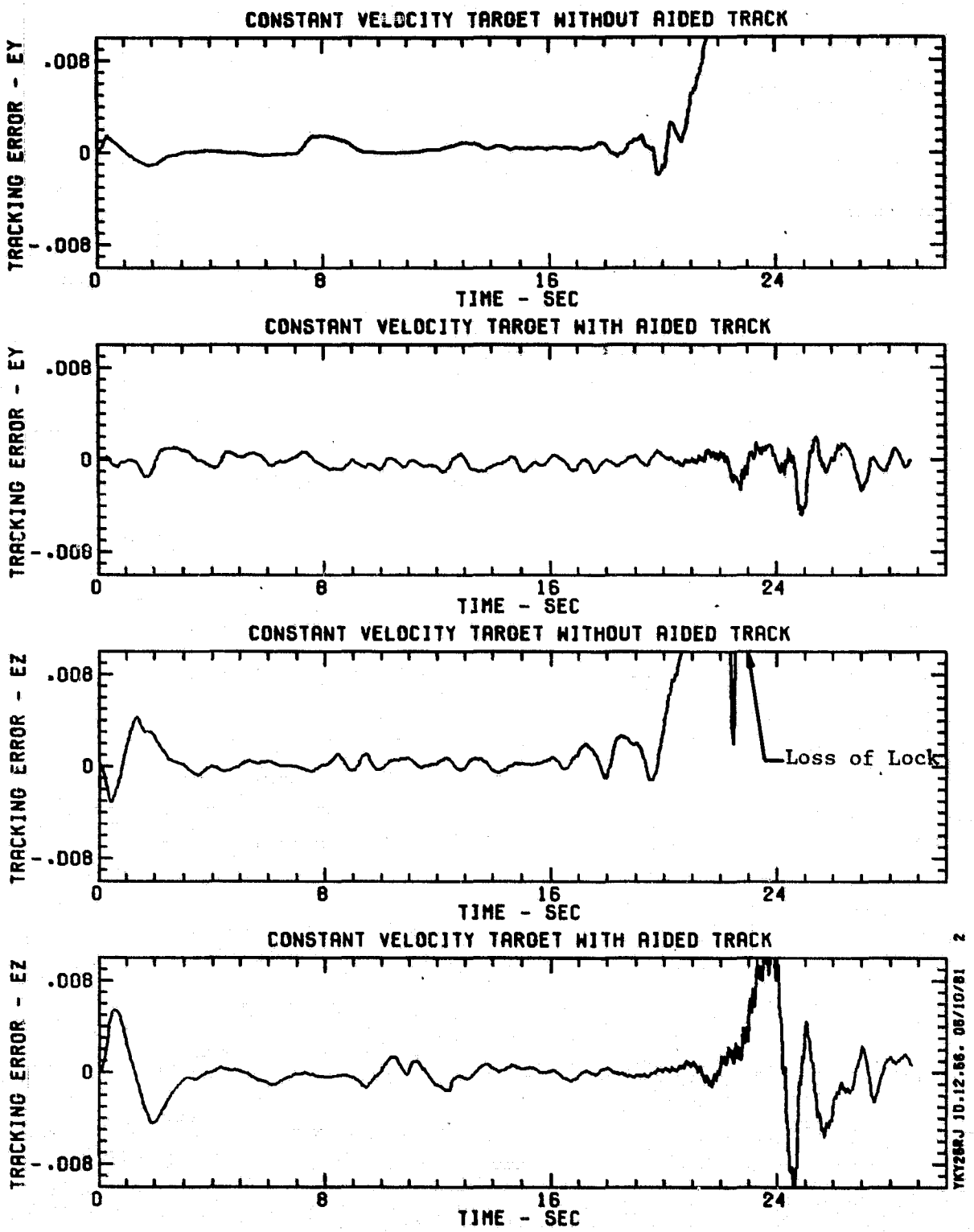


Figure 4. Tracking Error

DESIGN AND DEVELOPMENT OF A SIX DEGREE OF FREEDOM HAND CONTROLLER

M. L. King
Group Leader
SRMS

G. M. McKinnon
Manager
Research and Development

A. Lippay
Human Factors
Engineering

CAE ELECTRONICS LTD.

ABSTRACT

The design objectives of a six degree of freedom manual controller are discussed with emphasis on a space environment. Details covered in the discussion include problems associated with a zero g environment, the need to accommodate both 'shirt sleeve' and space suited astronauts, the combination of both manipulator operation and space craft flight control in a single device, and to accommodate restraints in space.

The lack of positive direction provided by a State-of-the-Art survey is discussed briefly as an introduction into the development work currently under way.

The initial work, consisting of a variable configuration device designed as a development tool in which rotational axes can be moved relative to one another, is described and its limitations discussed.

From work with the development tool two additional devices were developed for concept testing. Each device combines the need for good quality with its ability to achieve a wide range of adjustments.

The future work to be carried out towards an actual design proposal is described. The work so far indicates a trend to a particular type of handgrip for the space environment. When considering wider applications this type of grip may not recieve operator acceptance. Methods of mechanizing the same concept into more conventional forms are therefore discussed briefly.

The final section covers possible applications and the advantages which such a device could provide. Possible applications include the Shuttle Transportation System, its associated remote manipulators and appendages, and on-orbit or extended orbit space-craft.

INTRODUCTION

Co-ordinated manual control of multiple interactive devices is a common requirement. For a specific application, the ideal configuration of a controller depends on several factors: the relative importance of combined interdependent actions as opposed to sequential independent actions; the precision required; the working environment; other simultaneous tasks required of the operator. As a result, when one surveys available manual controls one finds myriad apparently unrelated devices, be it a steering wheel for a car, a control column for an aircraft or a set of levers for a back hoe operator. This paper blithely ignores the historical emphasis on a specific device for a specific application and addresses a more general problem, essentially responding to the question, "Can a six degree of freedom manual controller be designed to fit a general class of control problems?"

While the potential for general application was considered from the start, the design was guided by the requirements of certain specific tasks which imposed severe constraints. The stated objective was to provide single point control of six degrees of freedom on orbit in space, a non-dynamic zero 'g' environment. The tasks for such control indlude flight control of a craft carrying the operator and controller, control of manipulators attached to that craft, and the teleoperation of unmanned craft from a parent vehicle. An additional requirement is the operation of a large manipulator such as SRMS from a manned work station positioned at its outboard end.

The space application imposed two immediate design constraints: the controller must be suitable for use in a weightless environment; and it must be capable of being configured for operation with a heavily gloved hand. The protective glove worn by an astronaut performing extra vehicular activities severely limits manual dexterity and tactile feel. Further, the controller configuration has to be suitable for proportional control in all axes although an on-off or pulsed acceleration mode is required in some or all axes to achieve motion control of a spacecraft by the use of thrusters. Additional requirements of compactness, light weight and rugged mechanical design are generally beneficial in any application.

Ideally, the controller should enable the operator to command motions in any axis without crosscoupling while not inhibiting co-ordinated motion in any combination of axes.

The reason for the selection of six degrees of freedom is obvious in the case of motion control since six degrees are sufficient to determine the attitude and position of a rigid body. If co-ordinated, single hand control of six degrees of freedom is achievable, the order can be reduced by elimination of those axes for which motion is constrained. On orbit manoeuvers in space constitute a six degree of freedom control problem.

The chosen design should be adaptable to include additional controls to operate, for example, end effectors of manipulators.

Initial design activities included a literature search and a state-of-the-art survey which included discussions with many experts. The discussions, while lively and stimulating, produced a wide spectrum of opinion with few points of common agreement. There was general agreement, however, that a six degree of freedom controller was feasible and that a key factor in design is to ensure that the controller be compatible with the normative or mental model that an operator creates of his task. This implies that there should be spatial correspondence between the controller and the task, that is, up in the controller should correspond to up in the operators frame of reference and the forces applied to the controller should reflect the requirements of the task.

DESIGN CONSIDERATIONS

Various control modes and techniques were considered. The design selected evolved from selection of approaches which have been proven in practice. The selection process was subjective since it is difficult to compare results from previous studies due to the wide variation of applications, test conditions, quality of devices tested and personnel involved. The justification for some fundamental decisions is given here.

The first decision required was to select the mode, or modes of control required. Obvious candidates from a manipulation standpoint, were to use a replica controller or, alternatively, resolved rate control.

A replica control strategy involves using a scale model of the task in such a manner that manipulations of the model result in similar motions of the controlled device. In the case of a remote manipulator, this mode of control can provide excellent results. Problems arise from scaling, however, especially in the case of a large arm. A useable replica controller for the SRMS arm, for example, was not feasible since to model the 50 foot arm in the space available in the shuttle aft crew station would demand a large scale reduction requiring extreme precision in the master and with the effect that minor motions, nervousness or even the pulse of the operator result in significant control inputs. A twentieth scale replica controller has been used to provide excellent control of the RMS arm particularly in the case of the precise positioning required for docking the arm. A replica model has been implemented effectively using a one to one scale model for the hard suit arm which has been implemented and tested at several NASA laboratories including AMES and JPL. The replica approach has the advantage of providing excellent spatial correspondence and is adaptable to the use of force feedback. In some cases indexed position control has been used to aleviate the scaling requirements. However, this is achieved at the expense of spatial correspondence.

In the case of flight control of spacecraft, the concept of a replica type controller conflicts with the requirement for large unconstrained motions in all axes, and the requirement for a common fixed point of reference for the controller and the object being controlled.

For these reasons, the fact that the replica mode is task specific, and the excessive envelope requirements, the replica controller was rejected.

Two further alternatives were reviewed and rejected at this stage. One proposal was to use a proven, existing three or four axis controller with mode selection buttons so that one controller axis could be selected to control more than one axis in the task. This approach, while reducing the required mechanical complexity, and the design time for flight hardware, imposes constraints on co-ordinated motions and inhibits spatial correspondence. A second alternative, that of mounting a ring around a three axis translational controller in such a way that the ring could be rotated in the three rotational degrees of freedom, again made it possible to utilize existing hardware. However, the simplicity of single point control is lost.

The choice of six axis single point joy stick was considered the most general approach. While ingenuity would be required to achieve a feasible mechanical implementation, the resulting device could be used to implement resolved rate (or acceleration) or indexed position modes and did not impose any critical constraints in terms of hand position or spatial correspondence.

A second decision required was to select between isometric or force inputs and deflection inputs. Isometric controllers are rugged and easy to configure mechanically; however, they do not provide force feel or tactile feedback so that the operator can generate unwanted inputs and has a tendancy to overcontrol, especially when under stress. These shortcomings can be overcome with operator training and the relative merits of isometric versus displacement controls, as a general philosophy, are still a matter for debate.

For the six axes controller, the use of six isometric axes was rejected on the basis of evaluation of such a device built and tested at MIT. (See Figure 1) As a basic approach it was decided to use deflection in all six axes; however, the translational axes were designed so that the position input was measured indirectly as the force to deflect a spring. In this configuration, the translational deflections could be limited or locked out resulting in a device which uses deflection in the rotational axes and force for translation. The controller then, can be used either as a six axis deflection controller or in a "point and push" mode. The "point and push" mode has the advantage of simplicity in mechanical design while retaining good spatial correspondence features. Based on prototype evaluation, a selection will be made between the two modes.

The six degree controller was designed to provide adjustable force feel characteristics in all deflection axes. Force feedback was considered to be too difficult to implement at this stage. Force feedback or force cueing is, however, extremely important in many manipulator tasks and the possibilities of incorporating either force feedback or some form of force cueing is considered a high priority. True force feedback is achievable only in the case of replica or indexed position type of control; however, some form of force cueing in the resolved rate mode would be advantageous to provide information to indicate interference with stationary objects, external forces or high levels or applied force.

Based on the preceeding considerations two functionally similar, but physically different, devices were designed and built as described below.

BREADBOARD MODEL

A simple geometric breadboard model of a six degree of freedom controller was constructed to be used in evaluation. The model was designed to be adjustable in geometry, in particular permitting the six degrees to be about a single co-ordinate origin while allowing the yaw pivot to be displaced so that the yaw axis could be set either to the centre of the hand or to the wrist pivot point. The unit included light centering and breakout forces and position transducers; however, the inertia forces were large compared to the force feel characteristics. The unit is shown in Figure 2.

The breadboard unit was constructed with the handgrip placed inside the pivot points with the intention that, in later models, the rotational axes could be placed inside the handgrip to provide equivalent motions.

The breadboard model could also be used with a variety of handgrips.

Tests using the breadboard model were carried out to compare a wrist yaw pivot to a single point of origin in the hand centre. The null position of the hand was also evaluated and a novel handgrip evaluated which permits use with a gloved hand. The breadboard will be used in continuing evaluations.

PROTOTYPE MODEL

A prototype model as shown in Figure 3, was designed based on the results of the breadboard evaluation to provide six degrees of freedom about a single pilot located at the centre of the hand grip. The mechanisms for rotational motion and the rotational transducers are mounted inside the handgrip. The handgrip support is mounted on a three axis linear position device. Breakouts, gradients and hard stop positions are adjustable in each axis.

A design of the handgrip was based on several factors. First, to accommodate operation with a gloved hand, a substantial grip was required which conformed to a comfortable hand position. A raised portion was included to provide orientation. The use of a substantial handgrip permitted the rotational mechanisms and transducers to be packaged internally.

The handgrip mechanism, with its three rotational degrees of freedom was mounted on a three degree of freedom translation base. Two alternate configurations have been built, one allowing displacement inputs and the other isometric.

In the mechanical displacement configuration, position is sensed indirectly by means of a force transducer which detects the force applied to a linear spring. Since the breakout forces, force gradients and travels are adjustable, this configuration can also be used as an isometric device. One problem, typical of isometric controls, is that, in the presence of vibration, spurious control inputs can be generated due to the inertia and

mechanical dynamics of the controller. This effect can be eliminated in practice through the use of mechanical breakouts, electrical thresholds and careful consideration of controller structure and mounting configuration.

A second configuration in which the translational axes are purely isometric was also constructed. The configuration is mechanically simple and rugged and permits good feel characteristics in the rotational axes. Detents or breakouts can be included in the translational axes to prevent crosscoupling if necessary; the "point and push" mode of control will be investigated for a variety of applications.

TESTS

The prototype hand controller was designed to be flexible and adjustable both in terms of force feel and mode of control. Hardware interfaces have been designed which include adjustable threshold, independent gradient adjustment in each direction in each axis and adjustable saturation level.

To date, only 'non task related' tests have been carried out which have demonstrated the capability of generating single axis inputs as well as co-ordinated inputs in up to six axes.

In the immediate future more extensive laboratory testing is planned followed by the evaluation of the controller in various practical applications. The first of these will be as a control device for the Open Cherry Picker (OPC). Suited astronauts will use the controller to 'fly' the Large Amplitude Space Simulator (LASS) six degree of freedom cherry picker simulator. Following these tests the controller will be evaluated in the Hand Controller Development Facility at the Johnson Space Center (JSC). This facility provides simulated spacecraft flight control and can accept a wide variety of control devices for evaluating. Subject to availability it is also planned to carry out evaluation with the Manipulator Development Facility (MDF), also at JSC. This is a full scale working mock-up of the Shuttle Remote Manipulator System (SRMS) which is used for both development work and astronaut crew training. (Currently the system uses two separate hand controllers, one for rotation and the other for translation.)

CONCLUSIONS

A six degree of freedom prototype hand controller has been designed based on a review of existing designs and an assessment of current technology. The unit is flexible in that displacement control can be compared to isometric control in the translational axes. The final assessment depends on rigorous testing using various modes of control in various realistic applications. By assessing the results of these tests it is anticipated that some fundamental principles concerning the use of six degree of freedom controllers can be established.

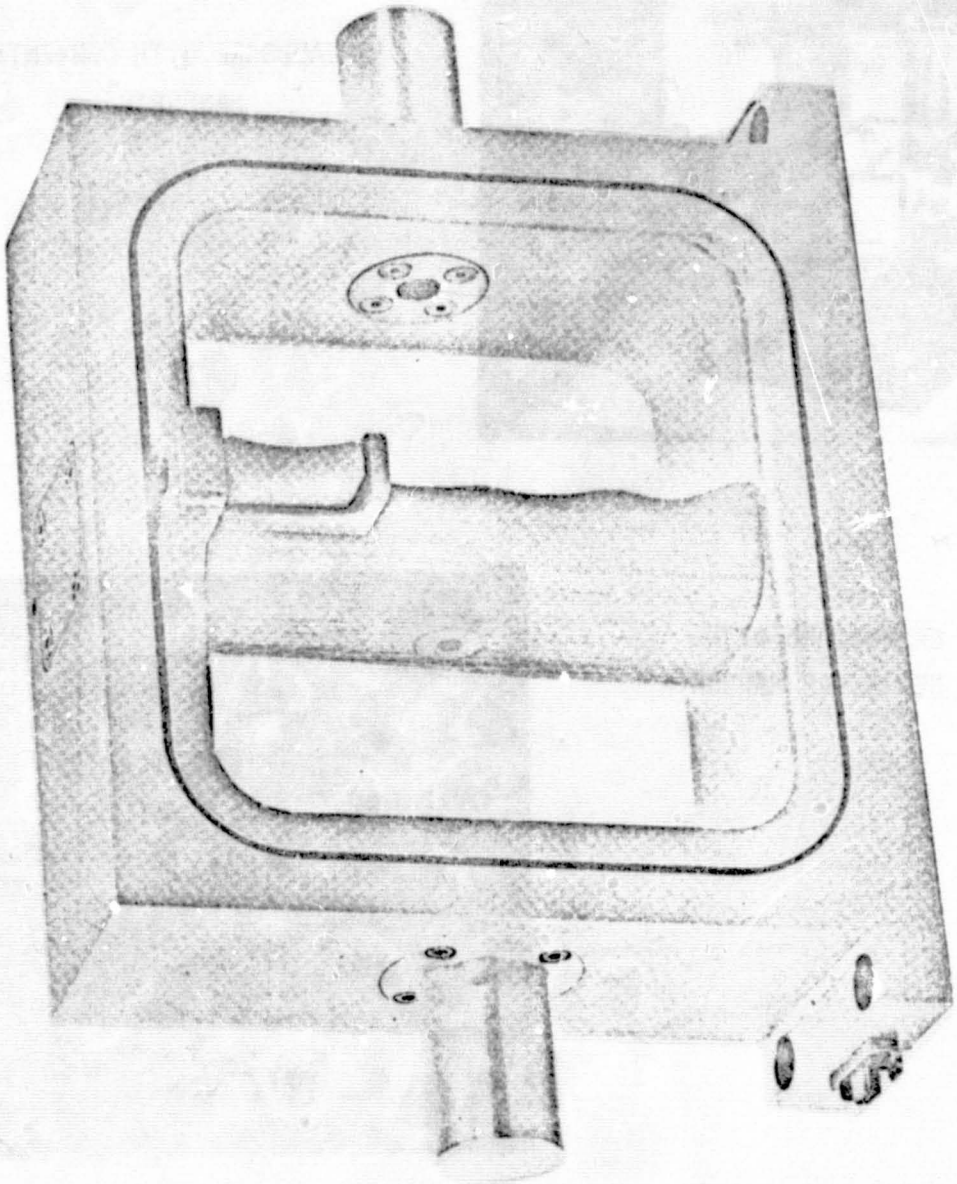
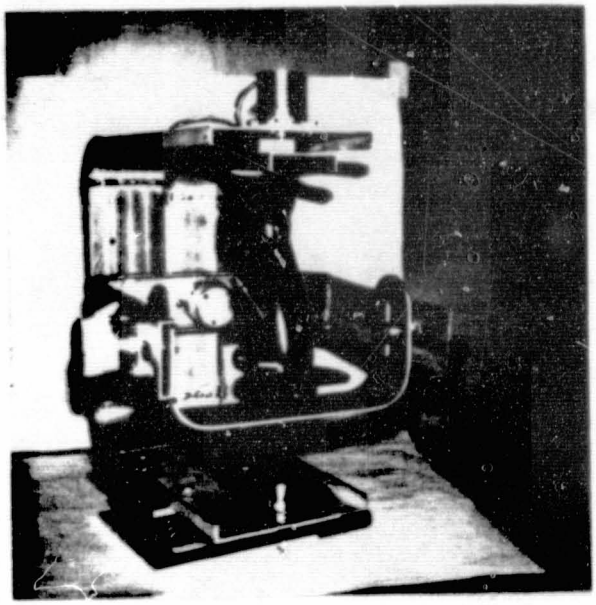


FIGURE 1 - MIT MK II ISOMETRIC HAND CONTROLLER



'BREADBOARD WITH CONVENTIONAL
HANDGRIP

BREADBOARD WITH
NEW STYLE HANDGRIP

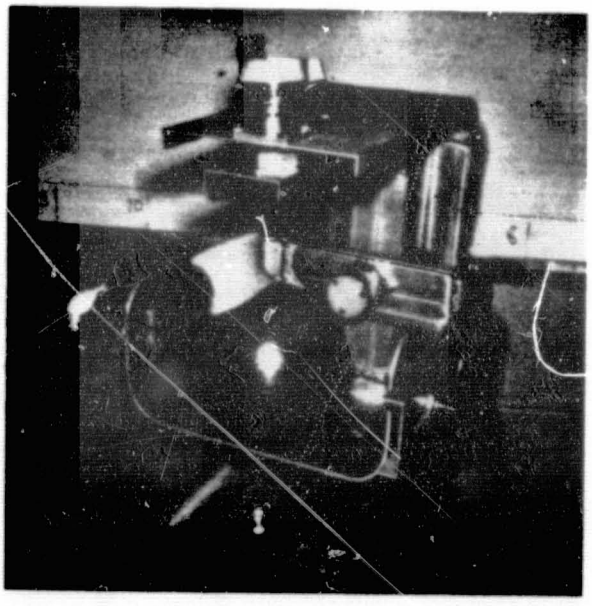
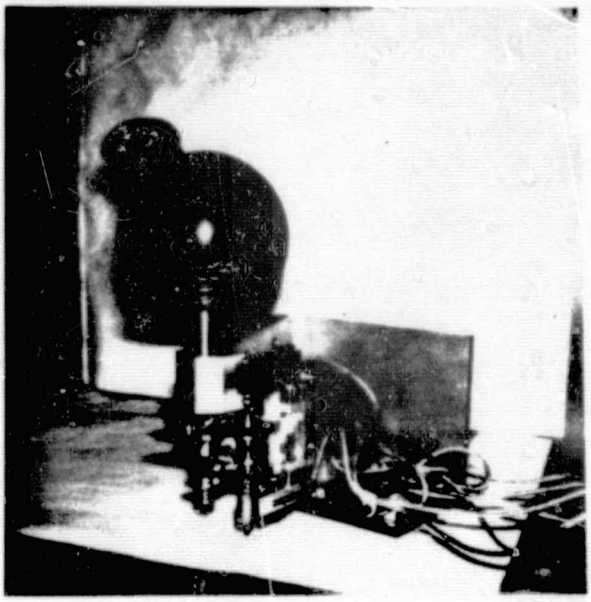


FIGURE 2 - BREADBOARD HAND-CONTROLLER



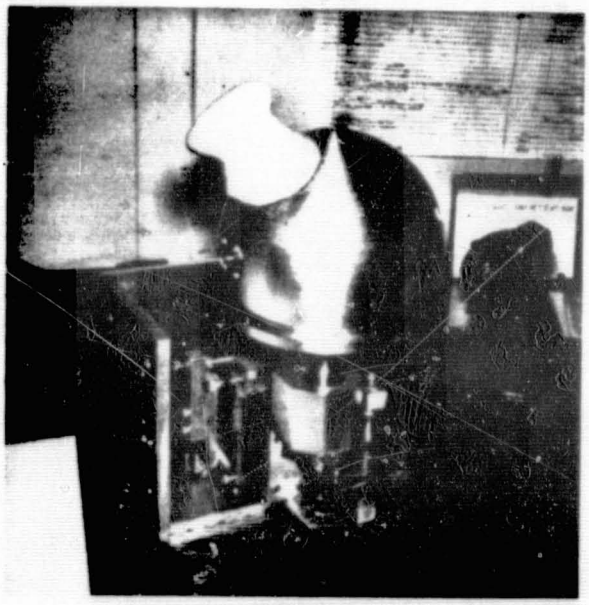


FIGURE 3 - PROTOTYPE HAND CONTROLLER

EXPERIMENTAL RESULTS WITH A SIX-DEGREE-OF-FREEDOM FORCE-REFLECTING HAND CONTROLLER

A.K. Bejczy and M. Handlykken*)
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

SUMMARY

The six-degree-of-freedom force-reflecting hand controller under current investigation at JPL is an isotonic joystick. Its hand grip is able to follow all the translational and orientational motions an operator's hand can comfortably make within a 30 cm cube. Each degree-of-freedom of this joystick can be backdriven by a motor commanded by the forces and torques sensed at the base of the hand of a remote manipulator. Thus, the operator can "feel" the task he is controlling when this joystick is connected to a remote manipulator through a computer. The use of this joystick for remote manipulator control can generalize the bilateral force-reflecting control of manipulators. Generalization means that the "master arm" function can be performed by this "universal" force-reflecting hand controller which is dissimilar to the "slave arm" both kinematically and dynamically. This paper briefly summarizes and evaluates a few preliminary control experiments performed by using this hand controller connected to a six-degree-of-freedom manipulator equipped with a six-dimensional force-torque sensor at the base of the manipulator end effector. The preliminary control experiments were aimed at the investigation of the human operators' ability to command and control forces in different directions by varying (i) the information conditions and (ii) the values of the feedforward and feedback command gains in the bilateral control loop. The main conclusions are: (i) a quantified graphic display of force-torque information can considerably enhance the operator's ability to perform a quantitatively sharp force-torque control, and (ii) there seems to be a task dependent optimal combination of the feedforward and feedback command gain values which provide a dynamically smooth and stable bilateral control performance.

I. INTRODUCTION

In the current bilateral force-reflecting master-slave manipulator systems widely and successfully employed in the nuclear industry the master and slave arms are in essence identical and interchangeable (Refs. 1-5). A limiting factor for broadening the application of the force-reflecting master-slave manipulator systems is the nature of the master arm. Typically, the present master arms are large and heavy, and require a large operating volume.

A pilot development system has been implemented at JPL recently. The system utilizes a six-degree-of-freedom force-reflecting hand controller as a master arm in combination with a slave arm which is totally dissimilar to the

*) On leave of absence from the Technical University of Norway, Trondheim, Norway.

hand controller both kinematically and dynamically. The development system is briefly described in Section II. The overall system is a kinesthetically coupled man-machine system. The input-output characteristics of the human hand play a key role in the bilateral control implementation which requires the use of a computer. In Section III control experiments are described aimed at evaluating the human operators' ability to control forces using this general purpose hand controller in a bilateral control mode under varying information and control conditions. The conclusions are summarized in Section IV.

II. EXPERIMENTAL SYSTEM

The main mechanical elements of the development system are shown in Figure 1. They are: a six-degree-of-freedom manipulator, a six-dimensional force-torque sensor mounted to the base of the end effector, and a six-degree-of freedom backdrivable hand controller. A computer which performs the coordinate transformations and closes the control loop between the hand controller and manipulator, as well as the sensor, drive and interface electronics are essential elements of the overall system.

The key mechanical element is the hand controller*) which acts here as a generalized master arm. It is essentially a backdrivable six-dimensional isotonic joystick which has been designed to conform to the motion range of an operator seated at a console. Its hand grip is able to follow all the translational and orientational motions that the operator's hand can comfortably make within a 30 cm (about 1 ft) cube work space. The hand controller mechanism is self-balanced, and can be mounted horizontally (as seen in Figure 1) or vertically. The self-balanced mechanism together with low backlash, low friction and low effective inertia at the hand grip render this hand controller a "transparent" interface between the human operator and a remote manipulator. More on the hand controller mechanism can be found in Reference 6.

The hand controller performs a dual function. First, it provides position and orientation commands to the manipulator. Second, it provides force and torque feedback to the operator's hand from the manipulator. This hand controller does not have any geometric and dynamic similarity to the manipulator it controls. In that sense it is a general purpose device: it can be interfaced to any manipulator through a computer. The computer reads the joint variables measured at both the hand controller and manipulator. Based on these measurements, real time computer algorithms establish the positional and orientational control relations between the hand controller and the manipulator. Likewise, real-time computer algorithms determine the motor torques needed to backdrive the hand controller joints as a function of the forces and torques sensed at the mechanical hand in order to provide a force-torque "feeling" to the operator's hand that parallels the force-torque "feeling" of the mechanical hand. The JPL/CURV manipulator, its kinematics, geometrical equations and control system together with the force-torque sensor integrated with it are described in detail in References 6-8.

---

*)The mechanism of the hand controller was designed by J.K. Salisbury, Jr., Design Division, Mechanical Engineering Department, Stanford University, Stanford, CA.

Figure 2 shows a simplified linear model of the bilateral control system dynamics referenced to one/one joint of the hand controller and manipulator. For simplicity, the geometric transformations are omitted from Figure 2. The overall performance of the bilateral control system is highly dependent upon the controller's ability of handling the interacting dynamics of the hand controller and manipulator. Note in Figure 2 that these two devices dynamically interact through the operator's hand. Note also in Figure 2 that the force-torque feedback to the operator's hand consists of three parts: (i) velocity damping, (ii) position error feedback, and (iii) feedback from the force-torque sensor. More on the bilateral control system analysis and synthesis can be found in References 9-10.

The simplified linear model shown in Figure 2 is only intended to illustrate two major points: (a) the general frame of the dynamic interaction between the manipulator, hand controller and the operator's hand, and (b) the meaning of the two control parameters, K_s and K_f, which were the key variables in the control experiments described below.

III. EXPERIMENTS

The purpose of the preliminary experiments was (1) to check out the overall performance and stability of the bilateral control system and (2) to evaluate the kinesthetic ability of the operator's hand to control prescribed forces in different directions when (i) the feedforward position scaling K_s and the force feedback gain K_f were changed and (ii) with or without using graphic display of force-torque information.

Four basic control experiments were performed:

1. Push down and hold 50N (~10 lb) force.
2. Push down and glide laterally with 50N (~10 lb) force.
3. Push forward, hold 50N (~10 lb) force, and zero out the lateral and down forces.
4. Push forward and down at the same time, hold 50N (~10 lb) force in each direction and zero out the lateral force.

In experiments 1 and 2 the task was set up so that the operator's wrist was free of lateral tension. In experiments 3 and 4 the task set-up required that the operator's hand be in lateral tension during the force control test. Note that the main force control action was (i) along the line of gravity field in experiments 1 and 2, (ii) perpendicular to the field of gravity in experiment 3, and (iii) with 45 degree tilt relative to the field of gravity in experiment 4. Note also that experiments 3 and 4 required the simultaneous control of all three (F_x, F_y and F_z) force components explicitly. In experiments 1 and 2 only one force component (F_x or F_z) control was required explicitly; the control of the remaining two (F_y, F_z or F_x, F_y) force components was only required implicitly.

Figures 3 through 8 show a few representative samples of the more than 300 experimental data curves generated so far. The unit value of the force feedback gain (K_f = 1) was nearly equal to 5N (~1 lb). The unit value of the feedforward gain (K_s = 1) signifies a one to one correspondence between the hand controller and manipulator displacements; the value K_s = 0.5 means that a 10 cm hand controller displacement causes only a 5 cm manipulator displacement.

The labels V and G at the performance curves in Figures 3 through 8 are related to two different information conditions. For the V curves, the operators had only visual feedback from the task scene together with the manually perceivable force feedback. For the G curves, the operators could observe a real-time color graphic bar display of the F_x, F_y and F_z forces acting on the mechanical hand in addition to the manually perceivable force feedback.

The data to some extent are hardware and software dependent and influenced by training, learning and other external conditions, and the total data base is quite narrow. Therefore, a detailed data evaluation is not yet possible. However, the data obtained so far allow a few general conclusions.

IV. CONCLUSIONS

1) Generalization of force-reflecting bilateral control of "master-slave" manipulators is feasible in the sense that the master arm does not have to be a kinematic and dynamic replica of the slave arm.

2) There is a trade-off between K_s (position feedforward scaling) and K_f (force feedback gain) parameter values: higher K_s requires lower K_f, or conversely, to obtain a dynamically smooth and stable performance.

3) There seems to be an optimal combination of the K_s and K_f parameter values. The optimal combination may be task dependent.

4) The operator's body posture, including the posture of his arm and hand holding the hand controller relative to his body, has a major influence on the dynamic performance of the overall control system.

5) A quantified graphic display of force-torque information considerably aids the operator in performing a quantitatively sharp force-torque control through a bilateral force-reflecting control system. Under certain gain conditions and without graphic display of force-torque data the system can become unstable (Figure 8).

6) The speed and direction at which contact is established between the mechanical hand and object have a major influence on the stability of task performance.

7) It is desirable to have a stiffer control coupling between the hand controller and manipulator.

8) Higher force feedback capability is desirable in the hand controller.

Acknowledgement

The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under NASA Contract NAS7-100.

References

1. R. Goertz, Manipulator Systems Development at ANL, Proceedings of the 12th Conference on Remote Systems Technology, ANS, November, 1964, pp. 117-136.

2. L. Galbiati and C. Mancini, A Compact and Flexible Servo System for Master-Slave Electronic Manipulator, Proceedings of the 12th Conference on Remote Systems Technology, ANS, November 1969, p. 73.

3. C.R. Flatau, Compact Servo Master-Slave Manipulator with Optimized Communications, Proceedings of the 17th Conference on Remote Systems Technology, ANS, November, 1969, pp. 154-164.

4. J. Vertut, Experience and Remarks on Manipulator Evaluation, Performance Evaluation of Programmable Robots and Manipulators, NBS Special Publication No. 459, Washington D.C., October 1975, pp. 97-112.

5. D.G. Jelatis, Characteristics and Evaluation of Master-Slave Manipulators, Performance Evaluation of Programmable Robots and Manipulators, NBS Special Publication No. 459, Washington D.C., October 1975, pp. 141-145.

6. A.K. Bejczy and J.K. Salisbury, Jr., Kinesthetic Coupling Between Operator and Remote Manipulator, Proceedings of the International Computer Technology Conference ASME Century 2, Volume 1, San Francisco, CA, August, 1980, pp. 197-211.

7. A.K. Bejczy, Allocation of Control Between Man and Computer in Remote Manipulation, Second Symposium on Theory and Practice of Robots and Manipulators, Proceedings, Eds. A. Morecki and K. Kedzior, Elsevier Scientific Publishing Co., Amsterdam-Oxford-New York, 1977, pp. 417-430.

8. A.K. Bejczy, Manipulation of Large Objects, Third Symposium on Theory and Practice of Robots and Manipulators, Proceedings, Eds. A. Morecki, G. Bianchi and K. Kedzior, Elsevier Scientific Publishing Co., Amsterdam-Oxford-New York, 1980, pp. 301-322.

9. M. Handlykken and T. Turner, Control System Analysis and Synthesis for a Six-Degree-of-Freedom Universal Force-Reflecting Hand Controller, Proceedings of the 19th IEEE Conference on Decision and Control, Albuquerque, NM, December, 1980, pp. 1197-1205.

10. M. Handlykken, Dynamic Stabilization Methods for Bilateral Control of Remote Manipulation, Proceedings of the ISMM Symposium on Mini-and-Micro-Computers in Measurement and Control, San Francisco, CA. May 20-22, 1981.

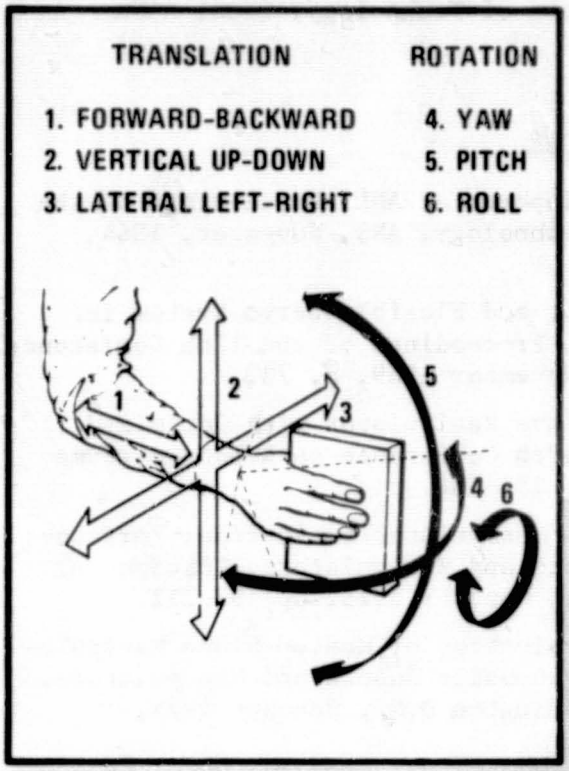
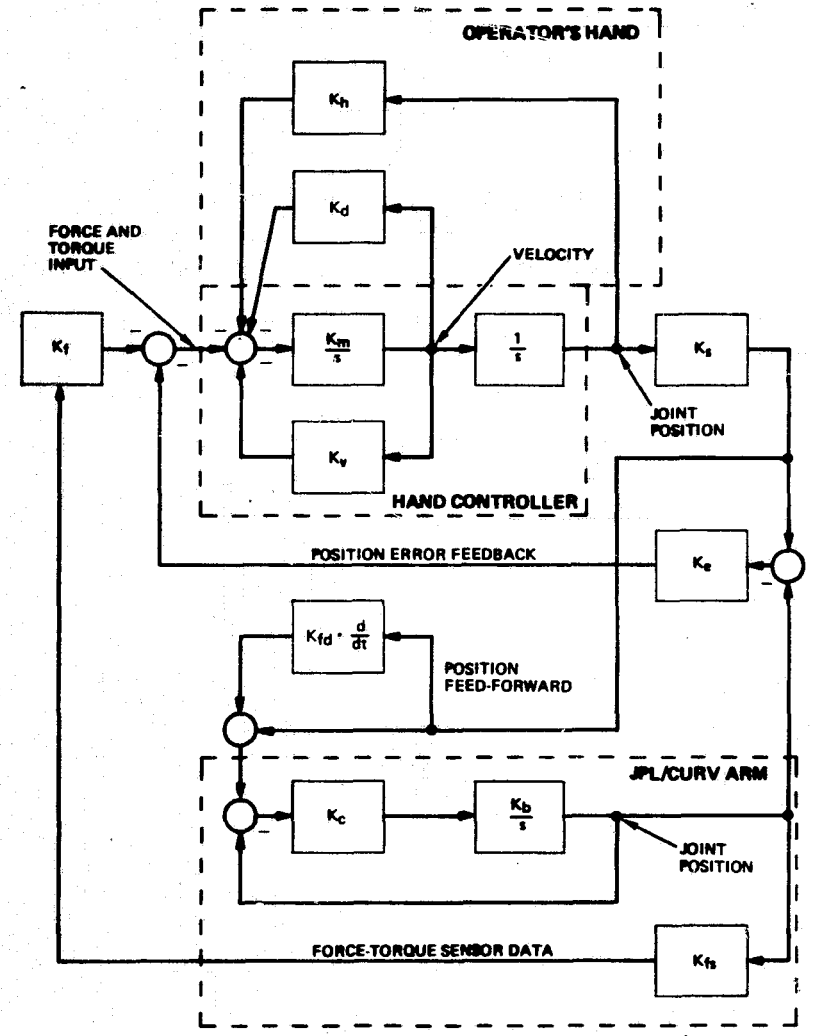


Figure 1. Overall Experimental System and Hand Controller Reference Frame



K_h: position correction in the operator's hand

K_d: damping in the operator's hand

K_m: aggregated hand controller constant formed by the product

$$K_m = K_t \frac{1}{I_f} K_a \text{ where}$$

K_t is the torque current constant

I_f is the hand controller inertia

K_a is the DAC/current constant

K_s: position feed-forward scaling

K_f: force feedback from force-torque sensor

K_v: velocity damping

K_e: position error feedback (to synchronize hand controller and manipulator motion)

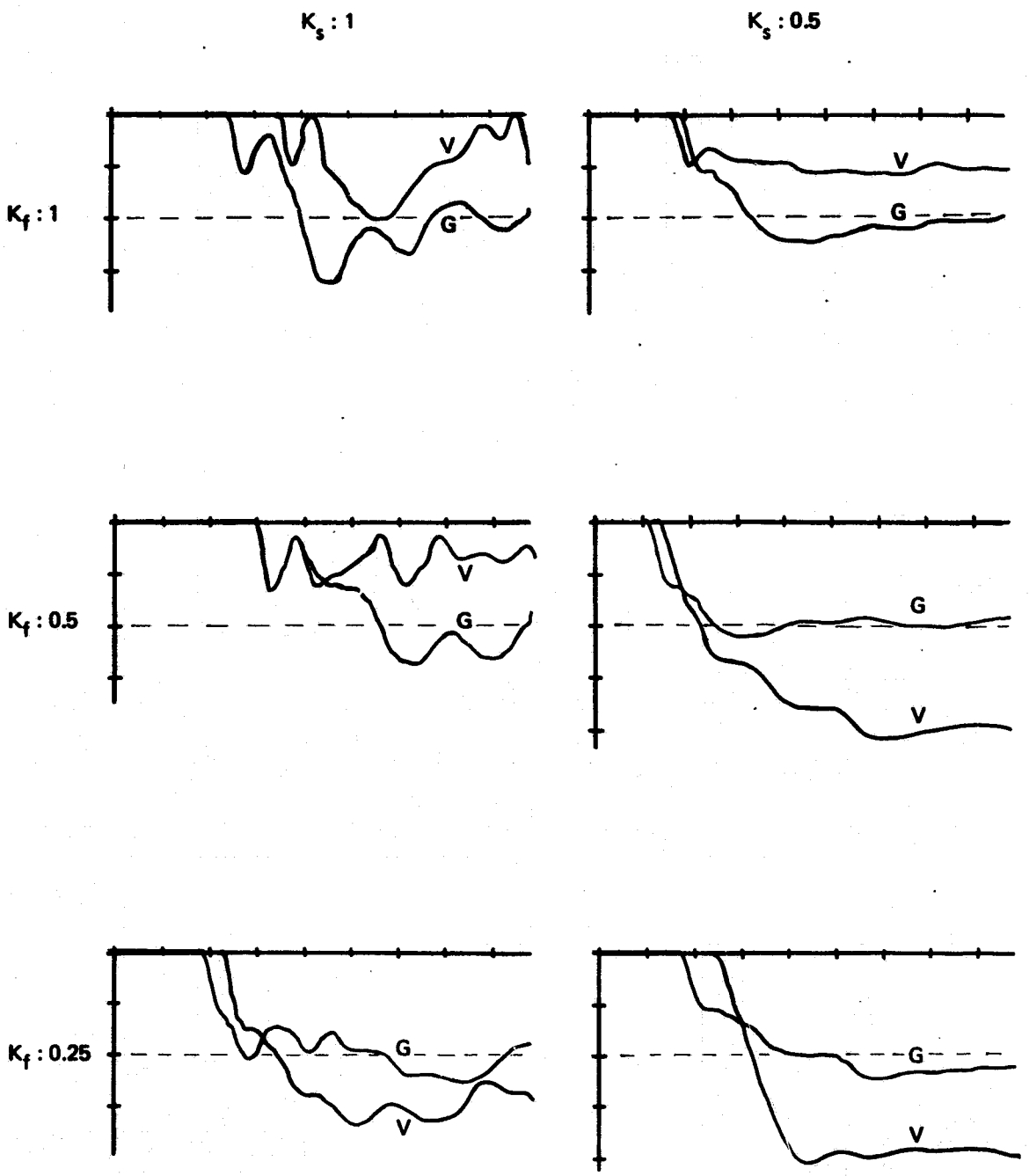
K_{fd}: position feed-forward to manipulator controller

K_c: manipulator controller gain (hardware implemented)

K_b: manipulator rate voltage constant (includes hydraulic system and amplifier)

K_{fs}: position/force constant (includes elasticity of both manipulator and object being contacted by end effector); it is zero when the end effector is not "forcing" any object.

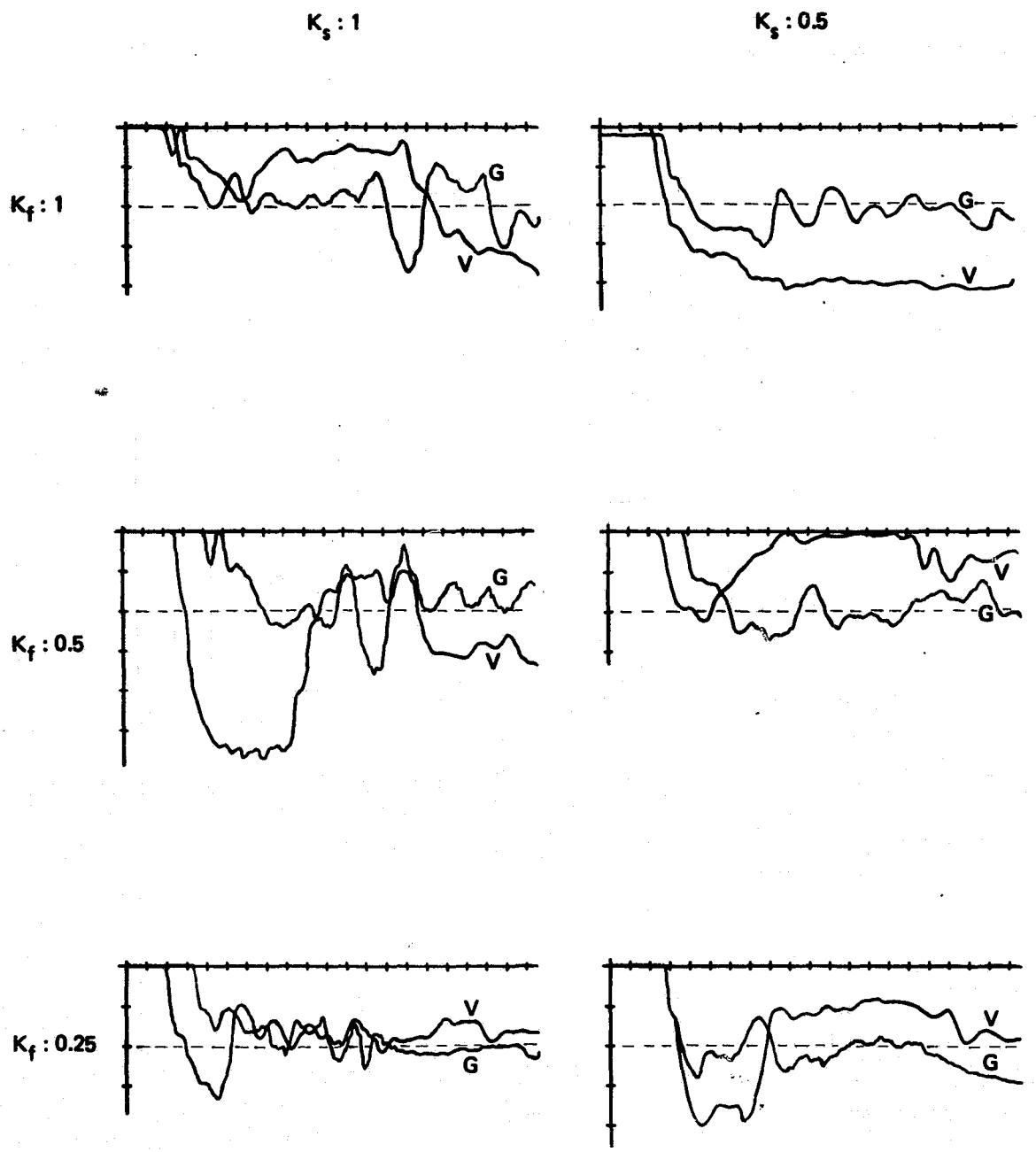
Figure 2. Simplified Bilateral Control Dynamics



VERTICAL AXES: FORCE, 25 N PER MARK
HORIZONTAL AXES: TIME, 1 SEC PER MARK

K_s : POSITION FORWARD GAIN
K_f : FORCE FEEDBACK GAIN
V : ONLY VISUAL SCENE OBSERVATION
G : ALSO GRAPHIC DISPLAY OF FORCE DATA

Figure 3. Push Down and Hold Experiments Data



VERTICAL AXES: FORCE, 25 N PER MARK
HORIZONTAL AXES: TIME, 1 SEC PER MARK

K_s : POSITION FORWARD GAIN
K_f : FORCE FEEDBACK GAIN
V : ONLY VISUAL SCENE OBSERVATION
G : ALSO GRAPHIC DISPLAY OF FORCE DATA

Figure 4. Push Down and Lateral Glide Experiments Data

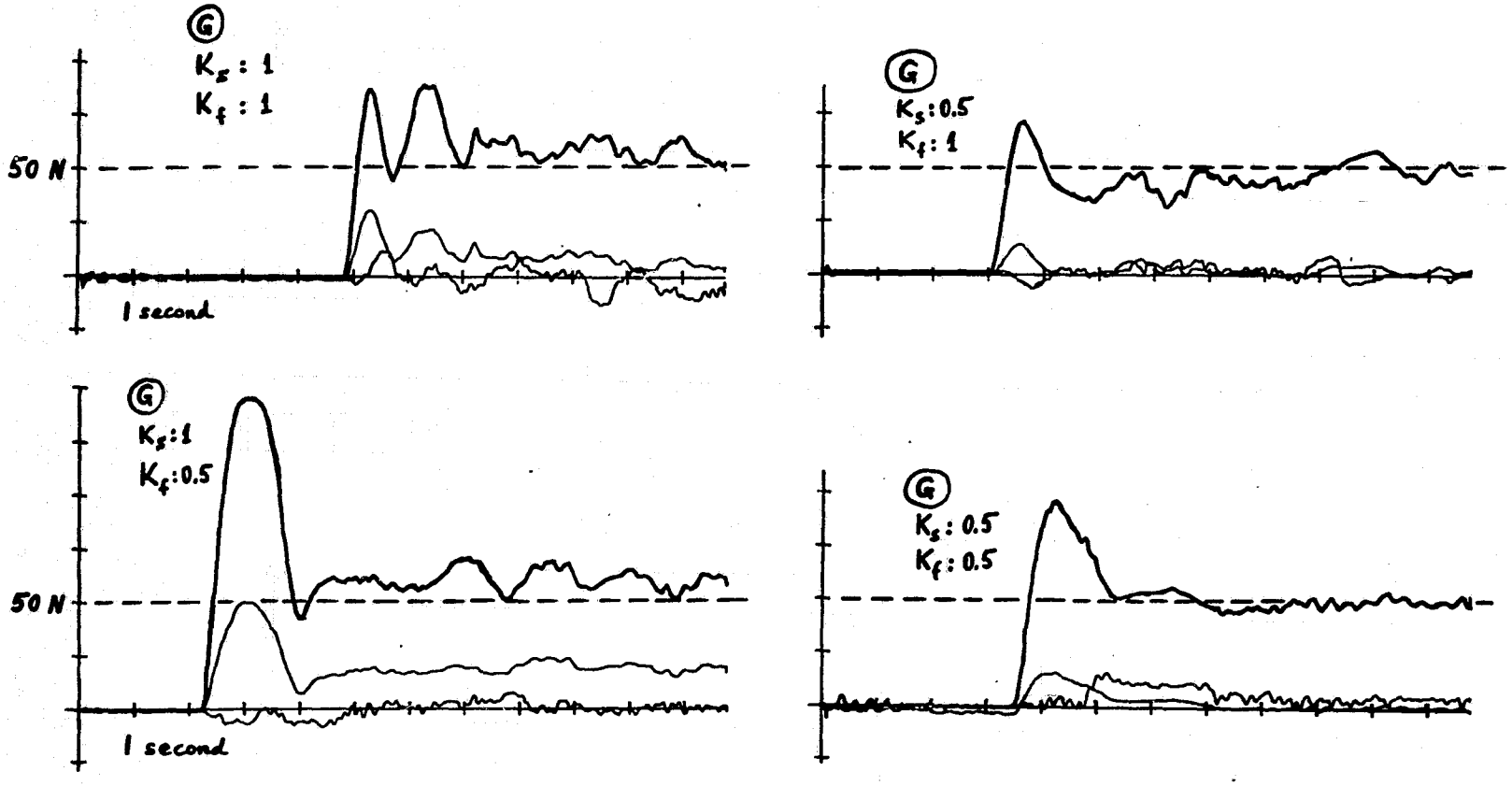


Figure 5. Push Forward, Hold Forward Force, and Zero Out Lateral and Down Forces, Using Also Graphic Display of Force Information

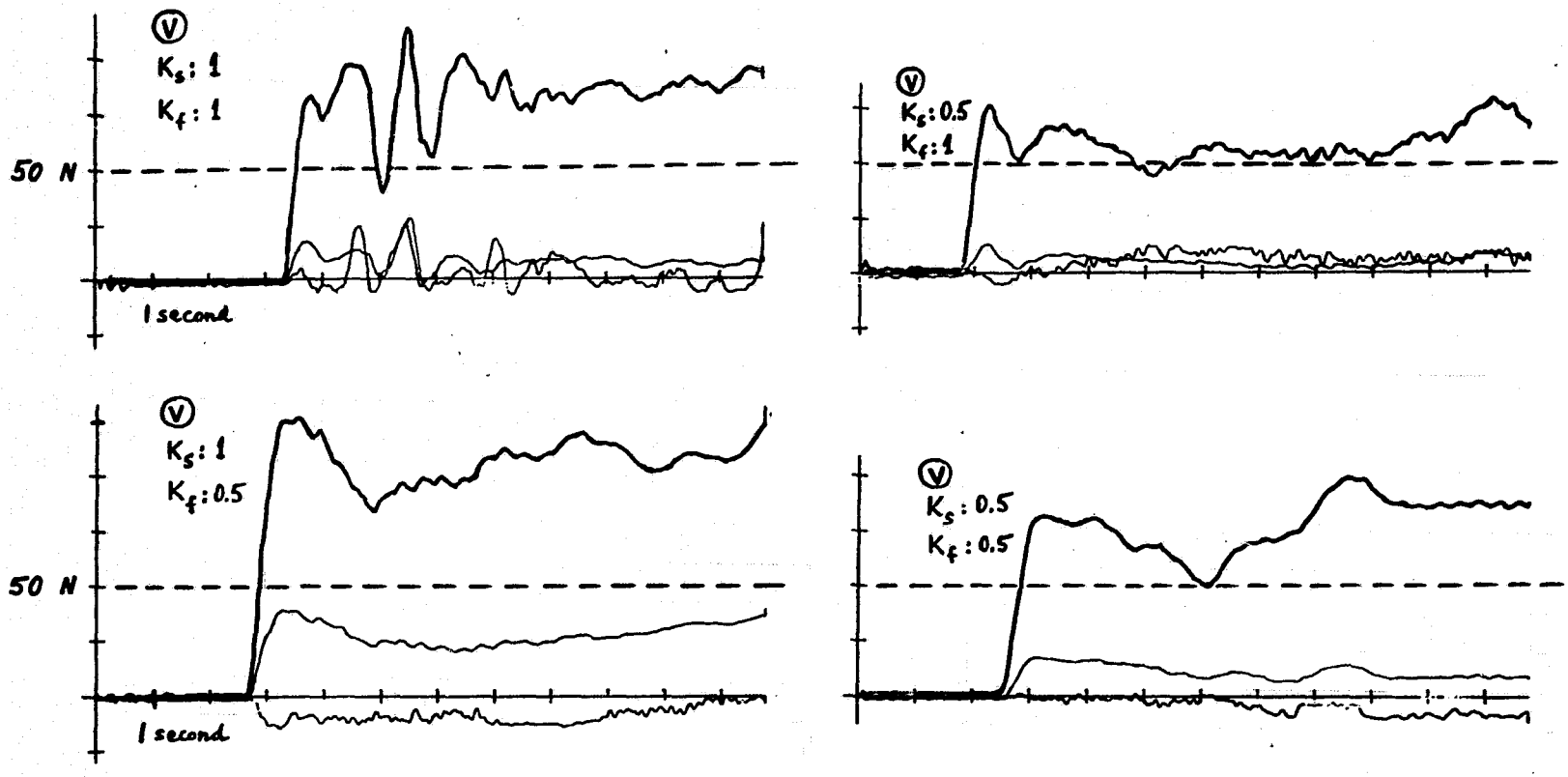


Figure 6. Push Forward, Hold Forward Force, and Zero Out Lateral and Down Forces, Without Graphic Display of Forces

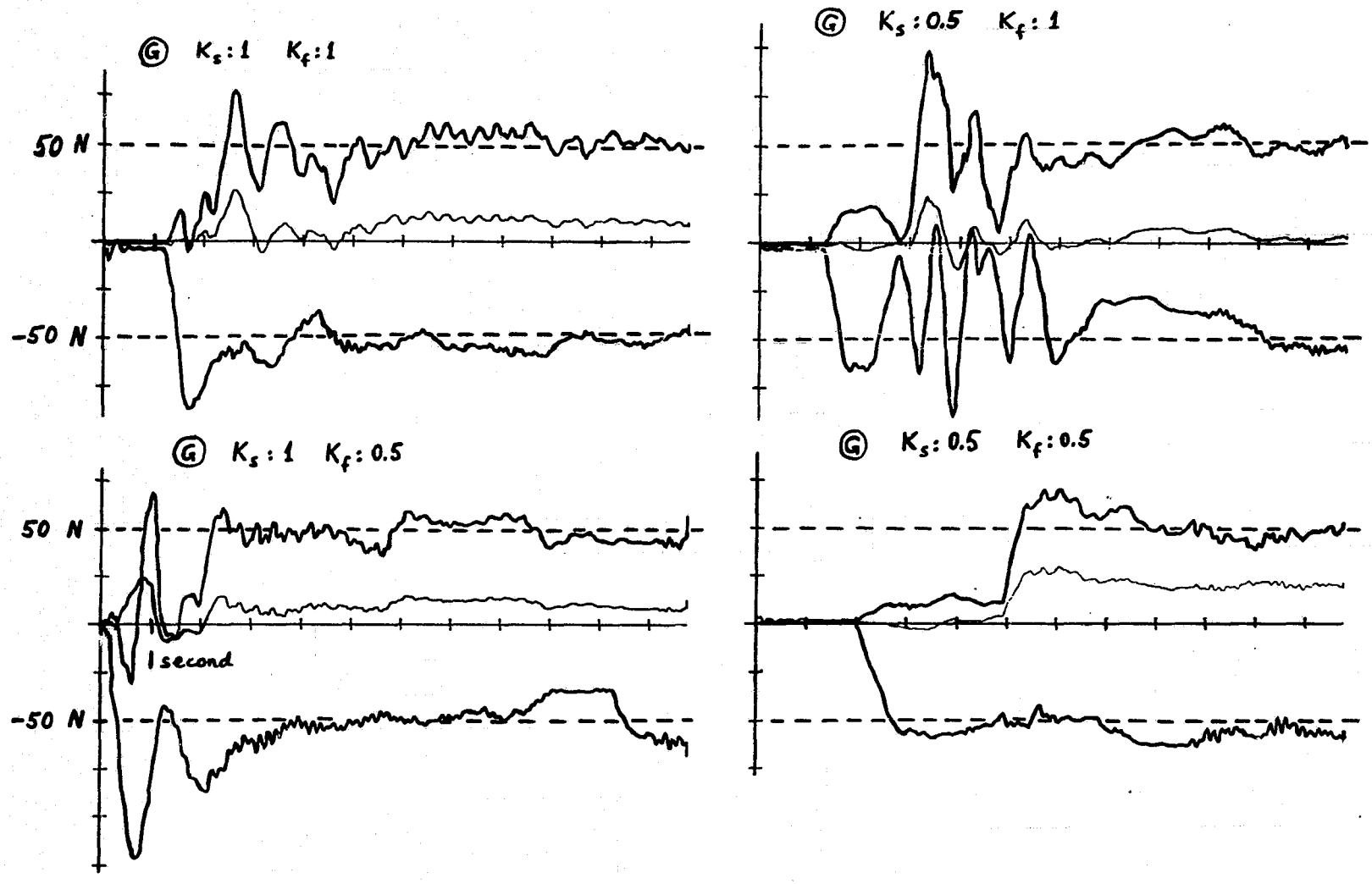


Figure 7. Push Forward and Down Simultaneously, Hold Force in Each Direction, and Zero Out Lateral Force, Using Also Graphic Display of Force Information

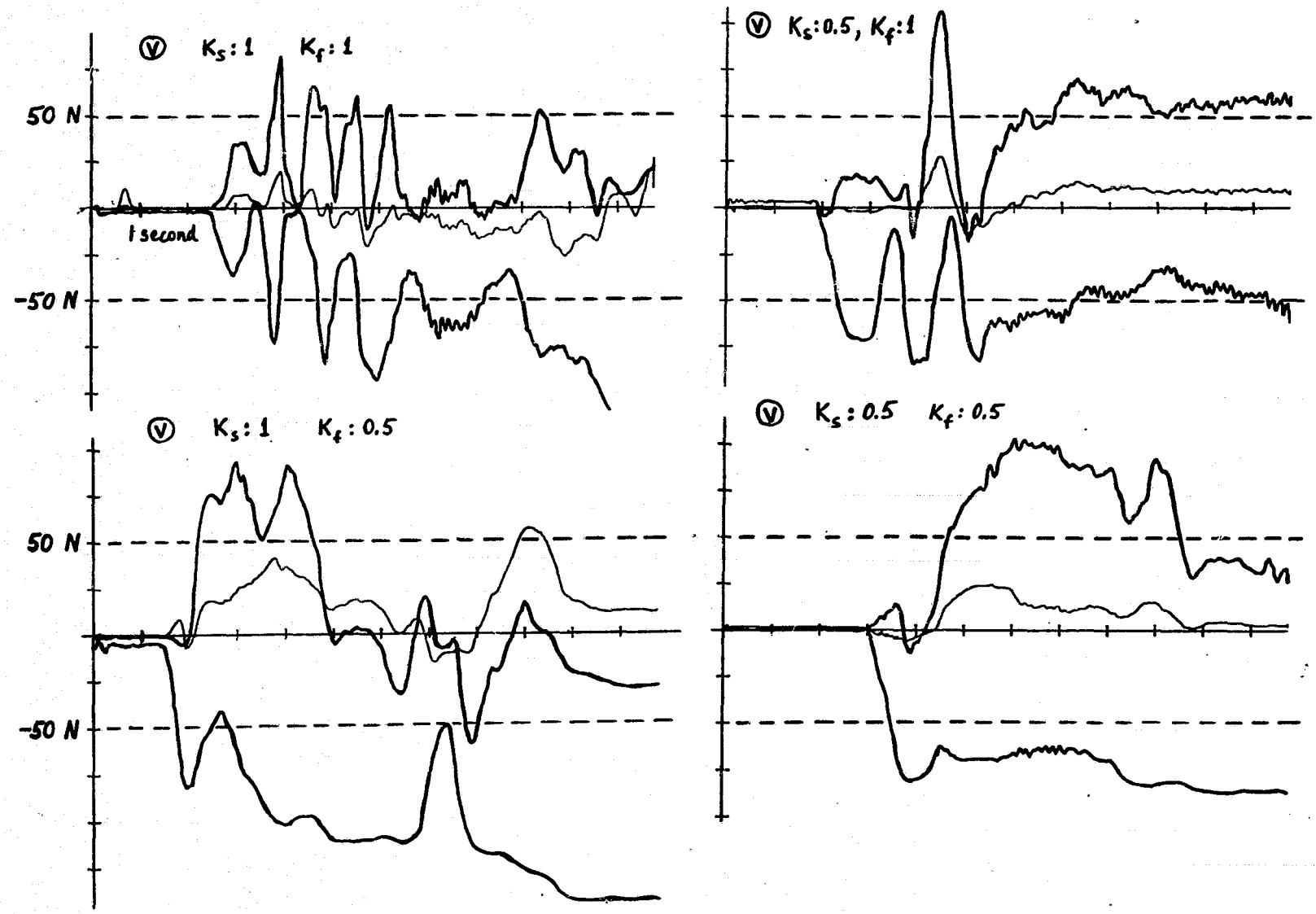


Figure 8. Push Forward and Down Simultaneously, Hold Force in Each Direction, and Zero Out Lateral Force, Without Graphic Display of Forces

LOAD COMPENSATING REACTIONS TO PERTURBATIONS AT WRIST JOINT IN NORMAL MAN

R. J. Jaeger, G. C. Agarwal, and G. L. Gottlieb

Bioengineering Program
University of Illinois at Chicago Circle, Chicago, IL 60680
and
Rush Medical College, Chicago, IL 60612

SUMMARY

The electromyographic responses to step torque loads were studied in flexors and extensors at the human wrist. Based on temporal bursting patterns and functional behavior, the response was divided into four temporal components. Two early components, the myotatic (30-60 ms) and late myotatic (60-120 ms) appear to be reflex responses. The third postmyotatic component (120-200 ms) appears to be a triggered reaction, preceeding the fourth, stabilizing component (200-400 ms). A comparison of responses at the wrist with similar data at the ankle provides the basis for a generalized classification of the responses in various muscles to torque step perturbations.

INTRODUCTION

The study of human motor control has recently seen a number of reports on the electromyographic (EMG) response of single muscles to highly contrived perturbations in carefully controlled laboratory situations. Since the work of Hammond (1956), many investigators have used the limb perturbation paradigm in an attempt to answer the question of how the human motor system responds to externally applied loads (Melvill Jones and Watt, 1971; Allum, 1975; Lee and Tatton, 1975; Marsden et al, 1976; Evarts and Granit, 1976; Crago et al, 1976; Gottlieb and Agarwal, 1976, 1979; Thomas et al, 1977; Hagbarth, et al, 1981; and others). As artificial as this paradigm is, it has remained the choice of many investigators. This is,

Presented at the Seventeenth Annual Conference on Manual Control, UCLA, June, 1981. This work was supported in part by National Science Foundation Grant ENG-7608754 and grants from the National Institutes of Health NS-00196 and NS-12877.

first, because of our lack of understanding and agreement at this most simple and controlled level, and second, because of the added complexities and technical problems involved in attempting to study the response of the motor system to perturbations in more complex paradigms.

One of the intriguing questions about the response of the human motor system to externally applied loads has been the extent to which it is achieved by reflex action versus voluntary action (see Weisendanger, 1978). Along with this question has been the as yet unresolved controversy concerning the neuroanatomical pathways mediating some of these responses. Views differ over the relative involvement of the cerebral cortex (e.g. Marsden et al, 1973). An alternative scheme links the responses to segmental mechanisms which mediate multiple bursts of afferent activity (Hagbarth et al, 1981). One purpose of the present study was to determine to what extent the EMG responses of wrist flexors and extensors to torque steps are mediated by reflex processes.

One might wish a hypothesis of the mechanisms for these responses to be applicable to the entire human motor system. Unfortunately, when the results from the more commonly studied muscles are compared, different muscles appear to have quite different responses to similar torque perturbations. This can give rise to different and even conflicting hypothesis of load compensation. For example, an early hypothesis suggested that the response of the human motor system to externally applied loads was to regulate length, and this hypothesis is commonly refered to as load compensation (Merton, 1953). A more recent hypothesis suggested that muscle stiffness was being regulated (Houk, 198Ø), that is, the system regulates both length and tension to maintain a certain level of stiffness.

Differing experimental observations and their explanatory hypothesis arise in part from the use of different muscles, or from slight differences in experimental paradigms among investigators. There are also the concomitant differences in terminology (e.g. FSR of Melvill Jones and Watt, 1971; M1-M2-M3 of Lee and Tatton; SL-ML-LL of O'Riain et al, 1979; myotatic-postmyotatic of Gottlieb and Agarwal, 198Øa).

Another purpose of the present study was to describe a data base of the EMG responses in wrist flexors and extensors that would be directly comparable to an existing data base on the ankle flexors and extensors (Gottlieb and Agarwal, 1979, 198Øa, b) using similar paradigms and subjects. This comparison is interesting in that it gives some indication of the possibility of devising a generalized classification of the response of the human motor system to externally applied loads, at least with regard to similar joints in the upper and lower extremity. Given the functional and anatomical differences between human upper and lower limbs, it is by no means obvious that a meaningful comparison would be possible, but such

proves to be the case. Such a general classification scheme is a useful step in formulating a unified understanding of the human motor responses to mechanical perturbations.

METHODS

All experiments were performed using 8 normal human subjects of both sexes, between the ages of 2Ø to 4Ø years. A subject sat comfortably in a chair next to the apparatus. The upper arm was abducted about 45 degrees, with the elbow flexed about 9Ø degrees. The forearm was aligned in the device, and either rested on a long foam pad or was supported by elbow and distal forearm supports. The wrist was positioned such that the axis through the head of the capitate about which flexion and extension movements occurred, coincided with the vertical axis of the shaft of a D.C. torque motor (Inertial Motors, Model Ø6-Ø24) mounted below the wrist. An aluminum bar was directly coupled to the shaft of the motor, and an adjustable flat paddle was bolted to the bar. The subject's hand, with fingers extended, was strapped flat to the paddle by three straps. Heavily padded mechanical stops limited the flexion and extension to about 75 degrees on either side of the relaxed neutral position. An oscilloscope display was provided with a reference target dot (finely focussed) and a movable response dot (diffusely focussed, about 3 times the width of the target dot) the position of which was continuously proportional to joint angle. This display assisted the subject in always using the same neutral rest position throughout the experiments. This apparatus restricted wrist motion to flexion and extension. No abducting and adducting movements could occur. A schematic of the experimental apparatus is given in figure 1.

Torque was measured by four semiconductor strain gauges (BLH Electronics, SR-4) bonded to the motor shaft (Perry and Lissner, 1955). Angular position was measured by a rotating-plate, capacitive transducer (Trans-Tek, model 6ØØ). The motor was driven by a D.C. power amplifier (Inland Controls, Model 2ØØB).

The EMG activity of the flexor carpi radialis (FCR) and extensor carpi radialis (ECR) was led off using redux creme column electrodes (Hewlett-Packard, model 1422ØA) with adhesive collars. The EMG signals were differentially amplified and band-pass filtered (1ØØ, 6Ø-1ØØØ Hz). This raw EMG signal was further amplified (15X), full-wave rectified, and finally passed through an averaging filter with a 7 ms averaging time (Gottlieb and Agarwal, 1970).

A digital computer generated the motor amplifier drive voltage. Typically, three levels of background torque or "bias" were used, with positive torques stretching the flexors. The computer digitized motor shaft angle and torque (at a rate of 25Ø samples per second) and FCR and ECR rectified and filtered EMGs (at a rate of 5ØØ samples per second) for later analysis.

The basic experimental paradigm was to deliver one second long steps of torque, randomized in amplitude, direction and inter-step interval to flex or extend the wrist. This was repeated for three different torque bias levels (Ø.Ø25, Ø, and -Ø.Ø25 kgm). The subject would be given one of four instructions on what to do when the torque step was received. These instructions were:

1. Do not react (DNR): Allow the motor to move the wrist without voluntary intervention.

2. React to target (RTT): React as quickly as possible to restore the wrist to its starting position.

3. React maximally (RMAX): React as quickly as possible to overcome it and move your wrist as forcefully as possible in the opposite direction, to the limits of the mechanical stops.

4. Assist (ASST): React by moving as rapidly as possible in the same direction to the limit of the mechanical stop.

Perturbation Experiments

Three series of perturbation experiments were performed. All three used random sequences of torque steps, evenly spaced between Ø.Ø7 and Ø.25 kgm in magnitude, presented at random intervals of from three to six seconds. In the first series, sets of thirty torque steps were presented with five step amplitudes randomly presented six times each. Step direction (flexion or extension) was always the same during one set, and this was known to the subject. Twelve sets were taken for each subject, with different combinations of instruction (RTT or DNR), bias torques, and step direction. This first series is analogous to a simple visual tracking experiment in the sense that subjects received a uni-directional perturbation which required a simple uni-directional response.

The second series of experiments presented one hundred torque steps (ten amplitudes, randomly presented ten times each). In this series, the direction of the torque step was also random, with five of the steps flexing the wrist and five extending it, and this was known to the subject. Six sets of data were taken for each subject, using two different instructions (RTT or DNR) at three values of bias torque. This second experiment was thus analogous to a choice visual tracking test with regard to the two possible directions of the perturbing torque step to which the subject selected an appropriate response.

In the third series of experiments, sets of thirty torque steps of five amplitudes were randomly presented six times. Step direction was always the same and this was known to the subject. A constant bias torque (Ø.Ø25 kgm, in the same direction as the torque step) was maintained in each set. A total of eight sets of data were collected for each subject, each with a different combination of instruction (RTT, DNR RMAX, or ASST) and direction of step/bias torque. Our subjects reported that the ASST instruction was the most difficult one to follow.

DATA ANALYSIS - PERTURBATION EXPERIMENTS

The data was analyzed, first, by averaging together the responses to torque steps of like amplitude and plotting the resulting ensemble averages versus time. The EMGs were then integrated over four selected response intervals: these were approximately 30-60, 60-120, 120-200 and 200-400 ms following torque step presentation. The first two intervals and the beginning of the third interval were chosen based on EMG bursting patterns found by inspection of the average plots. The boundary between intervals 3 and 4 and the end of interval 4 were chosen based on other considerations to be discussed later. The integrated EMG (IEMG) for each interval was corrected for the presence of background activity by subtraction of the mean IEMG level measured over a 50 ms interval before the torque step. There is some variation in these intervals between subjects which were determined by visual inspection of the data and adjusted in calculations of IEMG.

The IEMG for a given interval was then plotted versus the velocity of rotation, computed from the averaged angular velocity trace by digital differentiation 12-16 ms following the torque step, and a first-order, linear regression line was fitted. The slope of this line was taken as a measure of the gain of the reflex arc. For this reason, these plots are referred to as gain plots (see Gain of EMG Responses in discussion and Gottlieb and Agarwal, 1979, for further details).

The latencies of the first three EMG components were manually measured from individual records using an interactive graphics terminal. This was necessary because an unbiased measurement of latency cannot be obtained from averaged records.

REACTION TIME EXPERIMENTS

To compare the latencies of the perturbation-evoked EMG components against voluntary reaction latencies, another experimental paradigm was used to study visual and auditory voluntary reaction times, in a fourth series of experiments. Since no torque perturbations were used in these experiments the torque motor was used with velocity and position feedback circuits to provide a sensation of "springiness" rather than unimpeded rotation, which subjects reported as being helpful. For visual reaction times, the scope display was changed so that the computer controlled the position of the target dot which assumed one of three discrete positions on the screen: Center, and extreme right or left. The response dot moved along the same horizontal axis as the target dot.

The experimental paradigm allowed both simple and choice reaction times (SRT and CRT) to be measured. Initially, the target dot was at the center position. After a variable delay of three to five seconds it jumped left or right on the screen. The subject, instructed to track the target, chose the appropriate motion, a "choice" reaction. Once the target dot had moved and a response had been made, the target dot always returned to the center position,

again after a variable delay. Because the subject always knew the endpoint of the returning jump, this was a "simple" reaction.

Voluntary reaction times to auditory stimuli were studied in a similar paradigm. Here however, the target display was replaced with a voltage controlled oscillator, controlled by the computer to produce three distinct tones. The middle tone (≃ 33Ø Hz, "E") corresponded to the center neutral wrist position at rest. The low and high tones (≃ 15Ø and 5ØØ Hz, "D" and "B", respectively) corresponded to the extreme joint positions of flexion and extension, respectively. To begin the experiment, the middle tone was presented for about 1Ø seconds. After a random delay, the middle tone changed to either the low or high tone with the subject instructed to perform the appropriate wrist movement as quickly as possible. This was the choice reaction. Following this, the middle tone was again presented and the subject made a simple response to the center. Positioning the wrist in the center position was aided by a moderate amount of position feedback to the torque motor.

DATA REDUCTION - REACTION TIME EXPERIMENTS

The data collection and reduction was identical for both visual and auditory reaction times. Following a stimulus presentation, one second of data were collected at 25Ø samples per second of joint angle and EMGs from FCR and ECR. Latencies were measured from individual EMG records and tabulated as simple or choice responses.

RESULTS

Nature of EMG Responses

Typical responses to step torque perturbations are illustrated in figure 2. These are ensemble averages of joint angle and EMG from the stretched muscle (either ECR or FCR) for a single amplitude of torque step, in both flexion and extension, at three levels of bias torque. In both ECR and FCR, the EMG responses to a torque step were partitioned into four bursts. These we shall refer to as the myotatic responses (3Ø-6Ø ms), the late myotatic response (6Ø-12Ø ms), the postmyotatic response (12Ø-2ØØ ms), and the stabilizing response (2ØØ-4ØØ ms). The stabilizing response continued as long as the subject resisted the motor, but the first 2ØØ ms were taken as representative. In some records, each of the first three bursts is separated by short periods of silence, while in others the bursts overlap. The postmyotatic and stabilizing responses almost always appeared to merge. The rationale for choosing the intervals 3Ø-6Ø and 6Ø-12Ø ms can be seen from the bursting patterns of figure 2. Similarly, the choice of about 12Ø ms for the onset of the third interval is clear. The choice of 2ØØ and 4ØØ ms as interval boundaries was somewhat arbitrary, but this will be treated further in the discussion. Within these intervals there was always variability in the latency of each burst but rarely enough to obliterate the peaks and valleys of the averaged data. In any case, this categorization of responses relies,

not upon our ability to visually distinguish bursting patterns, but to make other distinctions of a more functional nature.

For most subjects the myotatic response is the smallest. It is not always present, and with the DNR instruction or a bias torque of opposite sign to the step it is frequently suppressed entirely. The late myotatic responses were larger and were present in every individual record. The postmyotatic and stabilizing responses, present in the lengthening muscle only with the RTT or RMAX instructions, were larger still. They were absent with the DNR instruction. With the ASST instruction, the postmyotatic and stabilizing responses appeared in the shortening (assisting) rather than the lengthening muscle.

GAIN OF EMG RESPONSES

In the first and second series of experiments, all four intervals showed a linear and monotonic increase of IEMG with the velocity of stretch (except for the last two intervals with the DNR instruction when no responses occur). The linear regression lines converged on the origin. There are other mechanical variables which correlate well with the velocity of stretch, such as the deflection angle and the amplitude of the torque step. The IEMG would show a similar behavior if plotted against one of these. It is because of this correlation between velocity and torque that the stabilizing responses (200-400 ms) is proportional to the velocity measure we use. At the ankle joint, the response in the interval 200-400 ms is proportional to the torque level in the motor being opposed by the subject (Gottlieb and Agarwal, 1980a). While this was not explicitly investigated at the wrist, we expect the same to be true.

The gain of the myotatic response depended heavily on the bias torque as shown in figure 3. This was most evident with the DNR instruction but also true for instructions requiring a reaction (RTT or RMAX). The behavior of ECR and FCR were similar to each other. The gain of the late myotatic response behaved much like the gain of the myotatic response in its dependence on bias, as shown in figure 4.

In distinction from the two earlier responses, the postmyotatic and stabilizing responses showed little dependence on bias, as shown in figure 5. The behavior of ECR and FCR were similar to each other for the two later components as well. The gain plots of figures 3 through 5 are representative of data seen in all 8 subjects.

The average gain for each subject and each IEMG interval were computed for experiments of series one and two. No systematic effect of a known versus an unknown direction of the torque step was seen. A standard t-test was used to test the hypothesis that the mean gain for series one (known direction) was the same as the mean gain for series two (unknown direction). In 19 of 24 cases (4 intervals by 6 subjects) there was no significant difference. ($p < 0.01$).

The dependence of the EMG gain in the four intervals on the full set of

instructions (RTT, DNR, RMAX, and ASST) from experiment 3 is shown in the gain plots of figure 6. Here, four regression lines appear on each response interval plot corresponding to the set of instructions. All regression lines for a stretched muscle converge on the origin and show a linear, monotonic increase in IEMG with stretch velocity with the exception of the RMAX line for the 200-400 ms interval. Note that the lines for the ASST instruction in the 120-200 and 200-400 ms intervals are taken from the assisting rather than the stretched muscle.

LATENCY IN VISUAL AND AUDITORY STEP TRACKING

The results for voluntary step tracking (series 4) are summarized in table 1. A series of t-tests were run within subjects to test the hypothesis that mean SRT was the same as mean CRT. This hypothesis was rejected in all but 2 of 12 tests at $p < 0.01$, although in these two cases SRT was still less than CRT. The fastest reaction time was auditory SRT, followed by visual SRT, visual CRT, and auditory CRT. Although there was a greater disparity between CRT and SRT in the auditory paradigm compared to the visual paradigm, the differences were significant in both. The most important point in these reaction time data is that the well known dichotomy (Welford, 1976) exists between voluntary simple and choice responses to visual as well as to auditory stimuli within our experimental setup.

EMG LATENCY - TORQUE STEPS

Latency measurements of postmyotatic responses (120-200 ms interval) were on the order of 120-150 ms, with no clear cut dichotomy between simple and choice reaction times. This was verified by testing whether the mean SRT was the same as the mean CRT using a t-test, as shown in table 2. Latencies of the 30-60 and 60-120 ms components remained stable over all experiments. The latencies of the 200-400 ms components could not be accurately measured, due to the overlap encountered with the previous interval.

The coefficients of variation for the postmyotatic, visual, and auditory latencies are summarized for simple and choice situations in table 3. Note that variability of latencies in choice situations is larger than in the simple situation for visual and auditory reaction times, but not for postmyotatic latencies.

DISCUSSION

Uniqueness of Each EMG Component

Each of the EMG components observed in this study displayed its own characteristic set of properties. The earliest response, occuring in the 30-60 ms interval, is the myotatic reflex. Based on latency considerations, it is equivalent to the M1 response described by Lee and Tatton (1975). It is analogous to the tendon tap response and thus largely mediated by the primary spindle afferents (Matthews, 1972). The wrist myotatic response is

both weak and variable when compared to the late myotatic response in the same muscles, or to the myotatic response in the soleus (Gottlieb and Agarwal, 1979). The gain of the myotatic response depends heavily on both instruction and bias. In most subjects, it could be entirely suppressed under appropriate combinations of those variables. Its latency was stable in all paradigms.

The second response (late myotatic), occurred in the 6Ø-12Ø ms interval. Based on latency considerations it is equivalent to the M2-M3 response of Lee and Tatton (1975). The afferent input mediating this response is uncertain, although it has been suggested on indirect evidence that the spindle may be the responsible receptor (Iles, 1977; Marsden et al, 1977; Chan et al, 1979). Recent data of Hagbarth et al (1981) suggests that multiple spindle discharges are responsible for the multiple reflex EMG responses observed in stretched muscles. However, the relationship between afferent stretch discharge and reflex EMG is not a simple one-to-one relationship. These authors noted that the EMG response often contained one less peak than the neural response. Other than the initial latency, the temporal correlation between peaks and valleys of the afferent activity and the EMG response is poor (see figs 6 and 7 of Hagbarth et al, 1981). The late myotatic response is much stronger than the myotatic response. Its gain could be modified by instruction or bias, but it was always present in every individual record. The latency of this response was stable over all data collected. There does not appear to be an equivalent late myotatic response in the soleus (Gottlieb and Agarwal, 198Øa), although there may be an analogous response in the biceps and triceps (Thomas et al, 1977).

The postmyotatic response was present only for the instructions involving a voluntary reaction by the subject. It is not a reflex (Gottlieb and Agarwal, 198Øa) and based on latency considerations, appears to be equivalent to what was described as voluntary activity by Lee and Tatton (1975).

Unlike its two predecessors, the postmyotatic response is not locked to the stretched muscle. Unlike visual or auditory reactions, the postmyotatic response shows no dichotomy between the simple and choice response situations. The gain curves of the postmyotatic response are like those of the myotatic and late myotatic responses for the RTT and RMAX instructions. They differ from the gain curves of the earlier responses in that they are insensitive to bias, and are zero when a DNR instruction is given. The postmyotatic response at the wrist appears to be quite similar to the postmyotatic activity occuring in the interval 1ØØ-2ØØ ms in the soleus (Gottlieb and Agarwal, 1980a). It remains to be seen whether analogous postmyotatic activity exists in the biceps or triceps.

Finally, the stabilizing EMG response in the interval 2ØØ-4ØØ ms occurs only for instructions requiring a sustained reaction. The IEMG for this component shows a monotonic increase with rate of stretch, except for the RMAX instruction, where a large offset is observed. The latency of this response could not be accurately determined.

Thus it appears that the EMG responses of wrist flexor and extensors may be partitioned into a minimum of four distinguishable intervals. Such a partitioning is compatible with the observed repetitive spindle bursting

described by Hagbarth et al (1981) but by itself such bursting cannot account for the observed functional differences between the different intervals.

One factor contributing to these differences may be the efferent fractionation of motor units in a similar experimental paradigm with monkeys (Bawa and Tatton, 1979; Tatton and Bawa, 1979). The differences may also reflect progressive recruitment and derecruitment of mechanisms from different levels of the motor contol hierarchy. From the present study no definite conclusion can be drawn.

REFLEX VERSUS VOLUNTARY RESPONSES

In considering various temporal segments of the EMG response, the question arises as to where to place the dividing line between reflex and voluntary responses. A short and stable latency, usually associated with reflex responses, would reflect primarily neural conduction time with minimal CNS processing delays. Certainly the myotatic response satisfies this criterion and may be considered a reflex response. Inability to voluntarily suppress a response is also a characteristic generally attributed to a reflex. This should not be confused with the ability to "voluntarily" modulate a reflex response (e.g. the Jendrassik maneuver), as illustrated by the dependence of the myotatic response on bias of the present study.

Similar observations apply to the late myotatic response. Although its latency is longer this response has all the characteristics of a reflex. The reflex arc associated with this response has been the subject of much controversy (e.g. Desmedt, 1979). No major distinctions other than latency can be drawn between these two components from the data presented here.

The postmyotatic response presents more of a problem. Its latency is longer than the preceding myotatic and late myotatic responses, on the order of commonly accepted values of kinaesthetic reaction times (approximately 12Ø ms, see Chernikoff and Taylor, 1952). Of course latency alone is not a sufficient classification criterion for any response. The postmyotatic response is present only with an instruction requiring a voluntary reaction by the subject. It does not depend on bias torque as do the previous responses, nor is it restricted to the stretched muscle. These are "voluntary" characteristics.

The postmyotatic response has other characteristics which differ from other responses we would commonly accept as voluntary. First, with the instruction RTT, an increasing IEMG with velocity of stretch is expected and observed. With the RMAX instruction a large, constant and velocity independent IMEG would be expected. In reality, the gain plots for the postmyotatic response with these two instructions are indistinguishable. The gain characteristics of the postmyotatic and stabilizing responses for the ASST instruction (measured in the assisting muscle) were not as consistent as their counterparts in the stretched muscle for the RTT and RMAX instructions.

A second feature of postmyotatic responses is that they show no dichotomy

between simple and choice reaction situations as do voluntary reactions to visual and auditory stimuli. It has been reported that at least for one type of kinaesthetic reaction time, no dichotomy between two and higher choice response tasks was found (Leonard, 1959).

In view of the above dilemma, we consider the postmyotatic response to be poorly described by either of the terms "reflex" or "voluntary". We call this a triggered response (Crago et al, 1976, Gottlieb and Agarwal, 1980a).

The exact point at which the truly voluntary response commences is open to question. Estimates from visual and auditory reaction times suggest a value on the order of 180-220 ms. In the present study, 200 ms was chosen. An exact determination is perhaps beyond the realm of neurophysiological definition.

Symmetry of Responses

Inspection of figure 2 shows that flexor and extensor responses to stretch are similar, possessing a high degree of symmetry. It was impossible to distinguish between FCR and ECR responses given the criteria of the present study. The symmetry in responses at the wrist is sharply contrasted by the asymmetry observed in ankle flexors and extensors (Gottlieb and Agarwal, 1979, 1980a). The response of the tibialis anterior is more like the responses of the FCR and ECR. The soleus differs, having a much stronger myotatic response and lacking the late myotatic response.

Known versus Unknown Perturbation Direction

No consistent effect of prior knowledge of torque step direction on the gain for any interval was seen. This is not in agreement with previous data on the effects of expectation on EMG responses to torque steps at the wrist (O'Riain et al, 1979)

General Characteristics of Responses to Torque Steps

Comparing the EMG responses and reaction time data for the wrist flexors and extensors of the present study with comparable data for ankle flexors and extensors leads to a scheme for classification of these responses. The schematic EMG responses to torque perturbations, shown in figure 7, form the basis for the following discussion. A schematic voluntary response to visual or auditory stimuli is also shown.

A period of silence immediately follows the torque step. This reflects neural conduction delays. The myotatic response (M1) is the earliest response, occurring at a latency comparable to that of the tendon tap response. At the wrist, it is enhanced by instructions requiring a reaction, while at the ankle, where the response is more pronounced in the soleus, it is less dependent on the instruction. Nevertheless, the myotatic response always: 1) shows a linear, monotonic IEMG increase with stretch velocity, 2) appears only

in the stretched muscle, 3) has a stable latency for all instructions, and 4) is strongly dependent on initial bias conditions.

The late myotatic response is not present in all muscles. At the wrist and tibialis anterior at the ankle, it is stronger than the preceding myotatic response. Like the myotatice response, it always: 1) shows a linear, monotonic IEMG increase with stretch velocity, 2) appears only in the stretched muscle, 3) has a stable latency for all instructions, and 4) shows a dependence on the initial bias conditions.

Whether this general scheme of classification will be applicable to other muscles remains to be seen. Its most notable feature is the de-emphasis of latency as the primary classification criterion. It appears to be useful for comparing the EMG responses to torque perturbation in the flexors and extensors of both the wrist and ankle. It proposes a progressive change in dominance, from an early dependence of the response almost entirely upon the stimulus to an eventual volitional dependence which is only dependent on the stimulus when the subject to chooses. The transition appears to take place in the 1ØØ to 2ØØ ms interval in which we observe the postmyotatic response. This transition interval appears to be the same at both the ankle and the wrist, indicating that conduction times between higher motor centers and the muscle play a small role in its determination.

REFERENCES

Allum, J.H.J. Responses to load disturbances in human shoulder muscles: the hypothesis that one component is a step test information signal. Experimental Brain Research 22:3Ø7-326, 1975.

Bawa, P. and Tatton, W.G. Motor unit responses in muscles stretched by imposed displacements of the monkey wrist. Experimental Brain Research 37:417-437, 1979.

Chan, C.W.Y., Jones, G.M. and Catchlove, R.F.H. The electromyographic response to limb displacement in man II: sensory origin. Electroencephalography and Clinical Neurophysiology 46: 182-188, 1979.

Chernikoff, R. and Taylor, F.V. Reaction time to kinesthetic stimulation resulting from sudden arm displacement. Journal of Experimental Psychology 43: 1-8, 1952.

Crago, P.E., Houk, J.C., and Hasan, Z. Regulatory actions of the human stretch reflex. Journal of Neurophysiology 39: 925-935, 1976.

Desmedt, J.E. (Editor) Cerebral Motor Control in Man: Long Loop Mechanisms. Karger, Basel, 1978

Dufresne, J.R., Soechting, J.F. and Terzuolo, C.A. Electromyographic response to pseudorandom torque distrubances of human forearm position. Neuroscience 3:1212-1226, 1978.

Evarts, E.V. and Granit, R. relations of reflexes and intended movements. Progress in Brain Research 44:1-11, 1976.

Gottlieb, G.L. and Agarwal, G.C. Filtering of electromyographic signals American Journal of Physical Medicine 49:142-146, 1970.

Gottlieb, G.L. and Agarwal, G.C. Dynamic relationship between isometric muscle tension and the electromyogram in man. Journal of Applied Physiology 30:345-351, 1971.

Gottlieb, G.L and Agarwal, G.C. Stretch and Hoffman reflexes during phasic voluntary contractions of the human soleus muscle. Electroencephalography and Clinical Neurophysiology 44:553-561, 1978.

Gottlieb, G.L. and Agarwal, G.C. Response to sudden torques about the ankle in man: myotatic reflex. Journal of Neurophysiology 42:91-106, 1979.

Gottlieb, G.L. and Agarwal, G.C. Response to sudden torques about the ankle in man: II - postmyotatic reactions. Journal of Neurophysiology 43:86-101, 1980a.

Gottlieb, G.L. and Agarwal, G.C. Response to sudden torques about the ankle in man III: suppression of stretch evoked responses during phasic contraction. Journal of Neurophysiology 44:233-246, 1980b.

Hagbarth, K.E., Hagglund, J.V., Wallin, E.U., and Young, R.R. Grouped spindle and electromyographic responses to abrupt wrist extension movements in man. Journal of Physiology (London) 312:81-96, 1981.

Hammond, P.H. The influence of prior instruction to the subject of an apparently involuntary neuromuscular movement. Journal of Physiology (London) 132:17p-180, 1956.

Houk, J.C. Regulation of stiffness by skeletomotor reflexes. Annual Review of Physiology 41:99-114, 1979.

Hufschmidt, H.J. and Hufschmidt, T. Agonist inhibition as the earliest sign of a sensory-motor reaction. Nature 154:607, 1954.

Iles, J.F. Responses in human pretibial muscles to sudden stretch and to nerve stimulation. Experimental Brain Research 30:451-470, 1977.

Jaeger, R.J., Gottlieb, G.L., and Agarwal, G.L. Afferent regulation of stretch-evoked myoelectric responses. (manuscript in preparation)

Lee, R.G. and Tatton, W.G. Motor responses to sudden limb displacements in primates with specific CNS lesions and in human patients with motor systems disorders. Canadian Journal of Neurological Sciences 2:285-293, 1975.

Leonard, J.A. Tactual choice reactions. Quarterly Journal of Experimental Psychology 11:76-83, 1959.

Marsden, C.D., Merton, P.A., and Morton, H.B. Is the human stretch refelex cortical rather than spinal? Lancet i:759-761, 1973.

Marsden, C.D., Merton, P.A., and Morton, H.B. Stretch reflex and servo action in a variety of human muscles. Journal of Physiology (London) 259:531-560, 1976.

Marsden, C.D., Merton, P.A. and Morton, H.B. The sensory mechanism of servo action in human muscle. Journal of Physiology (London) 265:521-535, 1977.

Matthews, P.C.B. Mammaliam Muscle Receptor sand Their Central Actions. 1972.

Melville Jones, G. and Watt, D.G.D. Observations on the control of stepping and hopping in man. Journal of Physiology (London) 219:709-727, 1971.

Merton, P.A. Speculations on the servo control of movement. in Wolstenholm, G.E.W. (Editor), Ciba Foundation Symposium: The Spinal Cord, Little Brown, Boston, pp 247-260, 1953.

O'Riain, M.D., Blair, R.D.G. and Murphy, J.T. Modifiability of the stretch reflex by prior information. Electromyography and Clincal Neurophysiology 19:57-63, 1979.

Perry, C.C. and Lissner, H.R. The Strain Gage Primer, McGraw Hill, New York, 1955.

Tatton, W.G. and Bawa P. Input-output properties of motor unit responses in muscles stretched by imposed displacements of the monkey wrist. Experimental Brain Research 37:439-457, 1979.

Thomas, J.S., Brown, J., and Lucier, G.E. Influences of task set on muscular responses to arm perturbations in normal subjects and parkinson patients. Experimental Neurology 55:618-628, 1977.

Welford, A.T. Skilled performance: perceptual and motor skills, Scott Foresman, Glenview, IL, 1976.

Wiesendanger, M. Comments on the problems of transcortical reflexes. Journal of Physiology (Paris) 74:325-330, 1978.

TABLE 1

COMPARISON OF SIMPLE AND CHOICE REACTION TIMES
FOR VISUAL AND AUDITORY STIMULI

VISUAL TRACKING

SUBJ	SRT			CRT				
	MEAN	SD	N	MEAN	SD	N	T	P
RJW	211	33	50	253	56	45	4.46	0.00
BKW	212	30	48	233	41	46	2.79	0.01
BMW	181	31	50	201	44	44	2.53	0.01
DVW	194	48	46	216	60	40	1.91	0.06
DMW	167	20	48	186	22	48	4.79	0.00
JGW	190	38	49	194	50	43	0.44	0.66

AUDITORY TRACKING

SUBJ	SRT			CRT				
	MEAN	SD	N	MEAN	SD	N	T	P
RJW	185	37	48	284	71	39	8.38	0.00
BKW	194	39	49	288	73	33	7.59	0.00
BMW	172	30	48	202	56	35	3.21	0.00
DVW	168	40	58	245	77	31	5.93	0.00
DMW	176	23	48	270	63	48	9.65	0.00
JGW	177	23	47	247	70	37	6.46	0.00

TABLE 2

LATENCY OF POSTMYOTATIC EMG COMPONENT
IN SIMPLE AND CHOICE REACTIONS
TO TORQUE PERTURBATIONS

SUBJ	SRT			CRT				
	MEAN	SD	N	MEAN	SD	N	T	P
RJW	131	17	126	128	15	217	-1.69	0.09
DVW	148	24	154	154	25	256	2.35	0.02
DHW	113	13	148	114	13	246	0.74	0.46
JGW	142	28	129	138	23	193	-1.42	0.16
BJW	122	15	141	125	16	189	1.73	0.08
REW	156	24	148	156	19	223	0.00	1.00

TABLE 3

COMPARISON OF COEFFICIENTS OF VARIATION
FOR POSTMYOTATIC LATENCIES AND VISUAL AND AUDITORY REACTION TIMES

SUBJECT	MR	EMR	PMR		AUDITORY		VISUAL	
			S	C	S	C	S	C
RJW	0.167	0.128	0.133	0.117	0.233	0.251	0.156	0.223
BHW	0.164	0.132	0.149	0.124	0.174	0.277	0.171	0.219
DVW	0.212	0.122	0.162	0.162	0.238	0.314	0.247	0.279
DHW	0.169	0.128	0.115	0.114	0.131	0.233	0.124	0.118
JGW	0.184	0.138	0.197	0.167	0.131	0.283	0.201	0.258

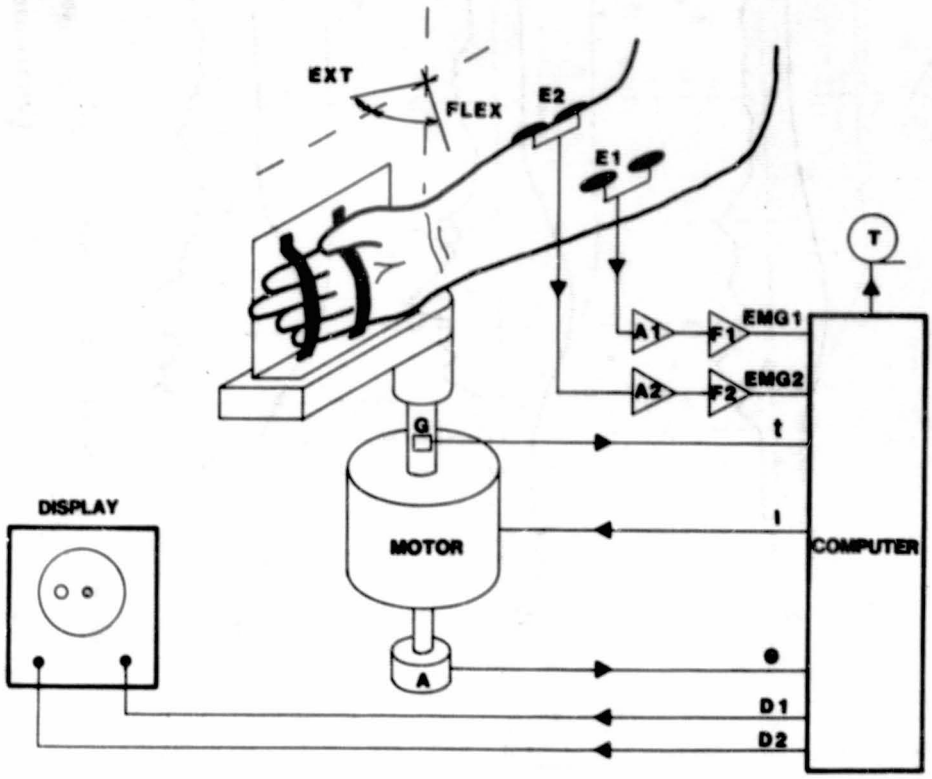


Figure 1: Schematic diagram of experimental apparatus consisting of a torque motor, whose current (I) was under the control of a computer. Joint angle (θ) is measured by a capacitive angle transducer (A), and torque (t) transmitted in the motor shaft is measured by strain gauges (G) bonded to the motor shaft. Electrodes (E1 and E2) led off the EMG through amplifiers (K) and filters (F). A visual display on the oscilloscope shows a reference target (D1) and joint position (D2). Data is stored off-line on digital tape (T).

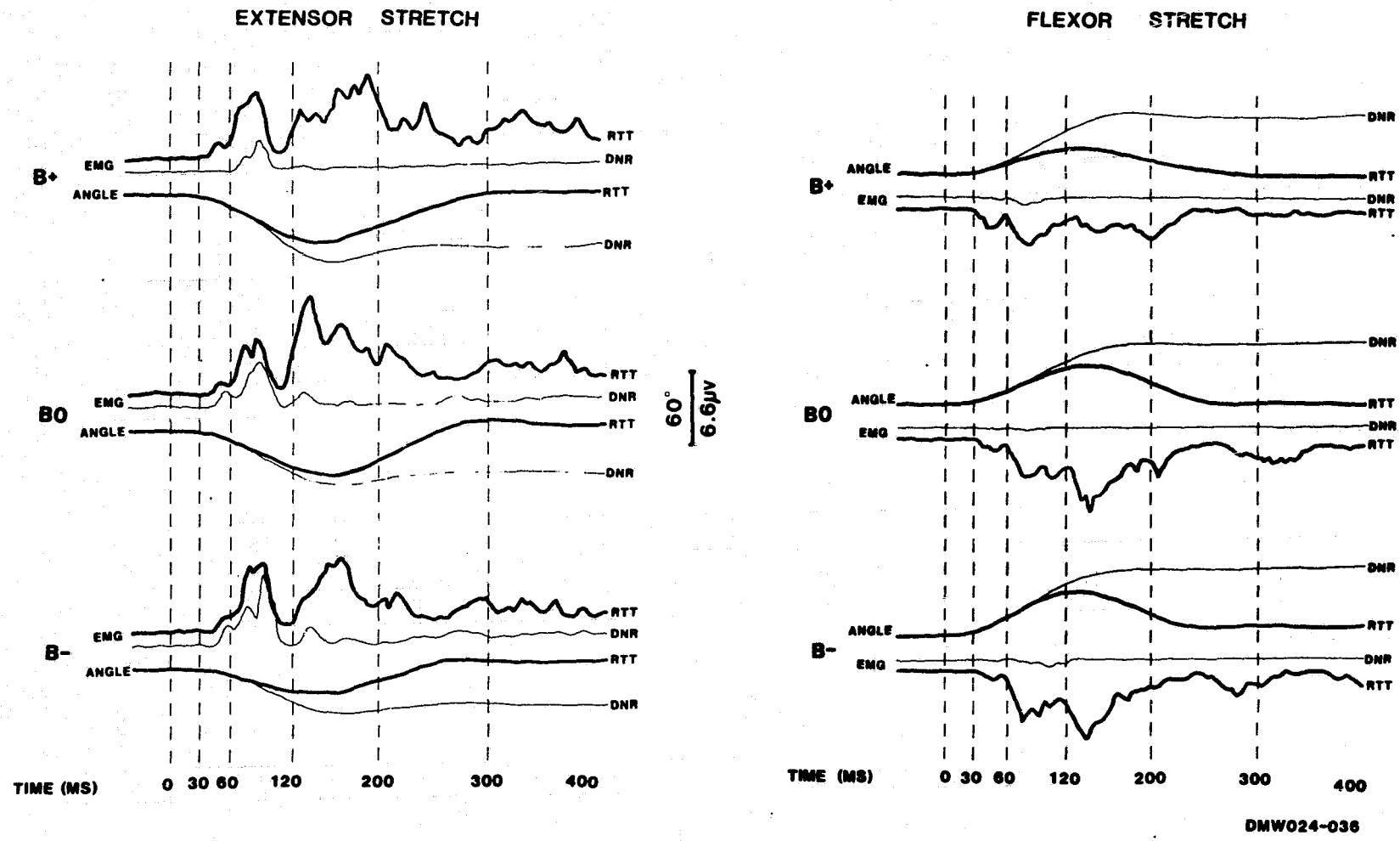


Figure 2: Joint angle and EMG responses to torque step in wrist extensor (A) and flexor (B) for three levels of bias. Instructions RTT (dark) and DNR (light) are superimposed. Torque steps are applied at time zero.

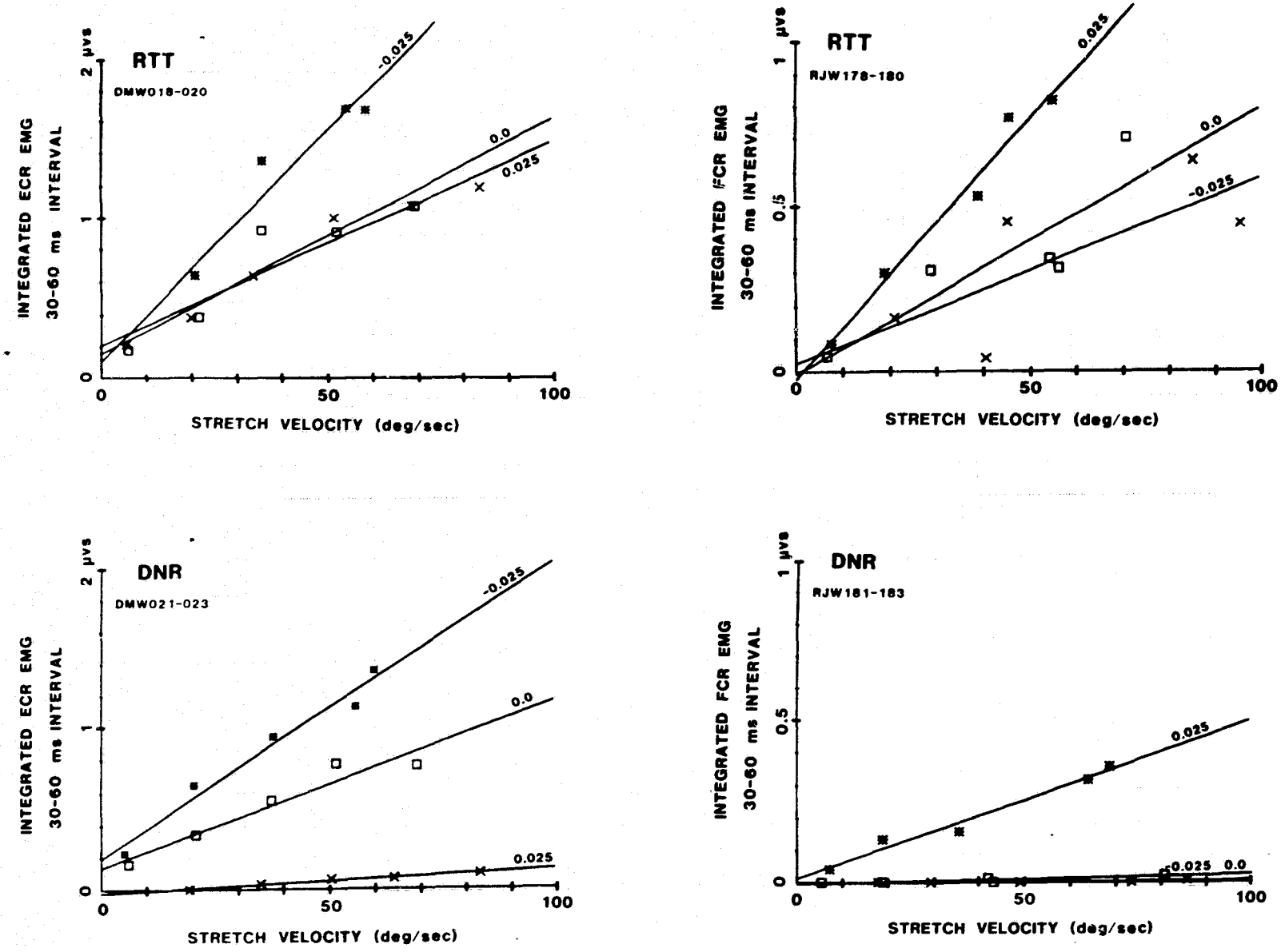


Figure 3: Gain plots for two subjects for myotatic response interval (3Ø-6Ø ms). Top plots for the RTT instruction, bottom plots for DNR instruction. Data from ECR for subject DMW and FCR for subject RJJ. Plot symbols: *-facilitory bias (bias and step have same sign), o-zero bias, x-inhibitory bias (bias and step have opposite sign). Regression lines labeled with values of bias in kgm.

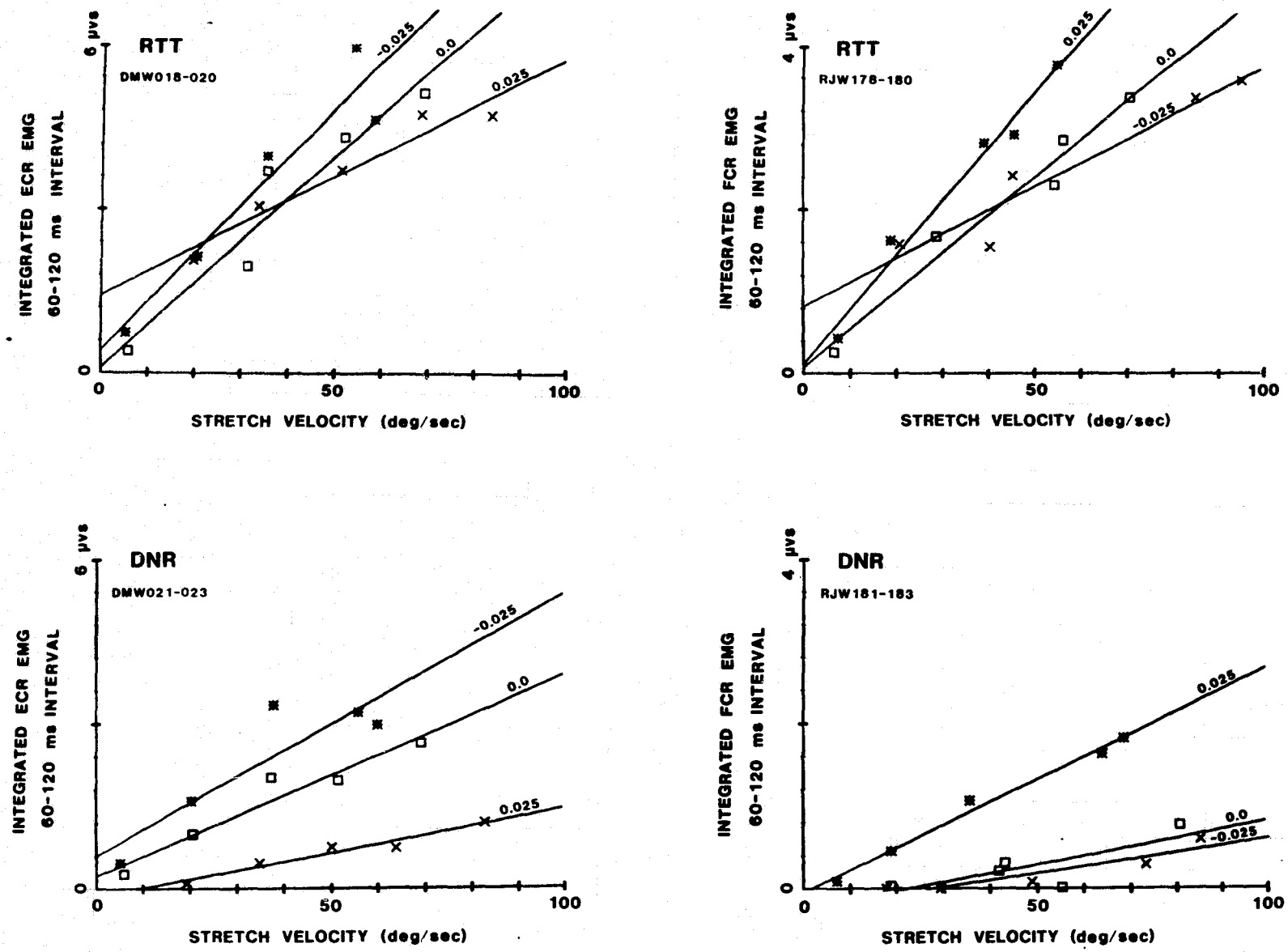


Figure 4: Gain plots for two subjects for late myotatic response (6Ø-12Ø ms). Top plots for RTT instruction, bottom plots for DNR instruction. Data from ECR for subject DMW and FCR for subject RJW. Plot symbols: *-facilitory bias (bias and step have same sign), o-zero bias, x-inhibitory bias (bias and step have opposite sign). Regression lines labeled with values of bias in kgm.

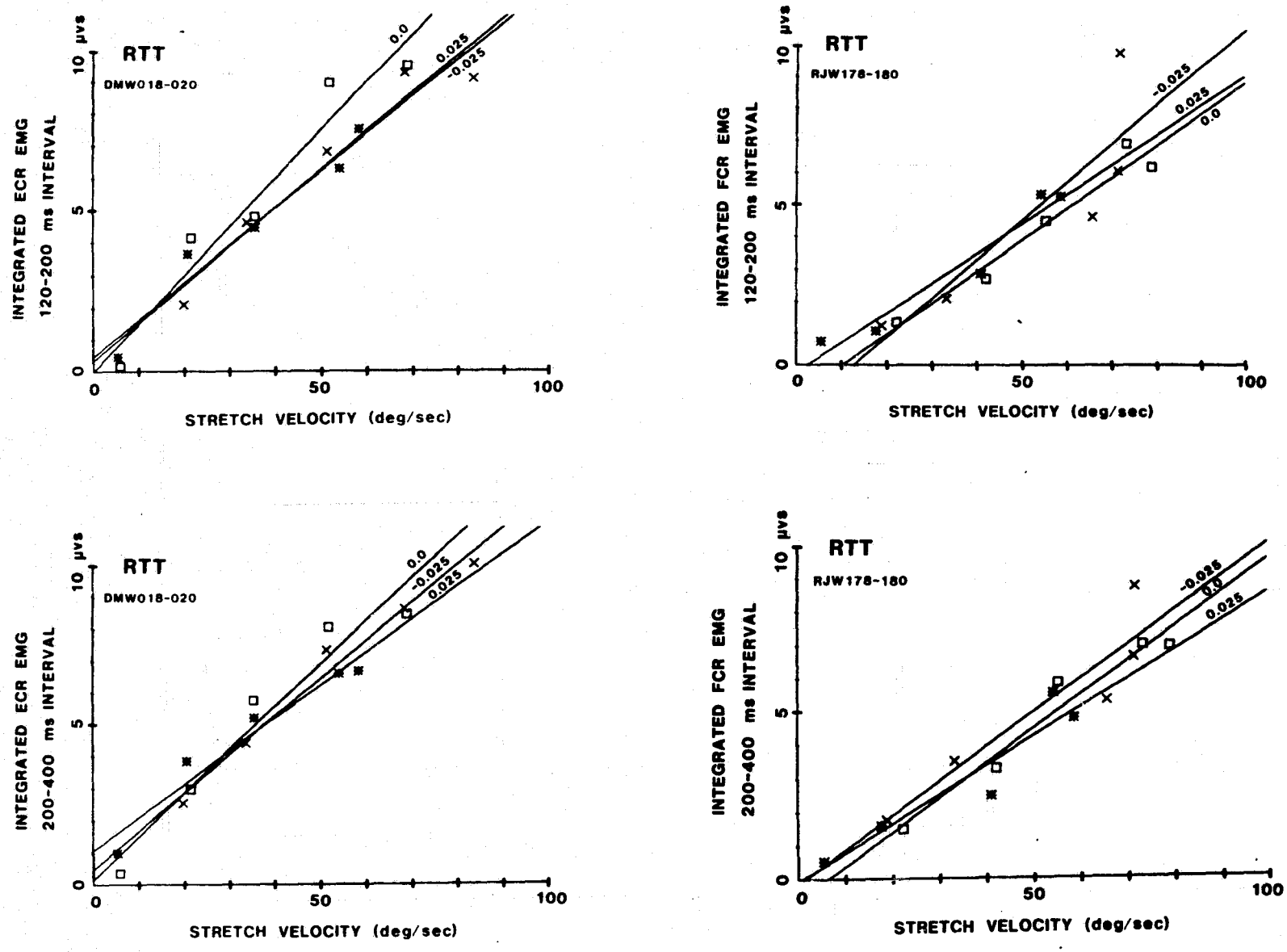


Figure 5: Gain plots for postmyotatic and stabilizing responses (12Ø-2ØØ and 2ØØ-4ØØ ms intervals respectively). Data from ECR of subject DMW and FCR of subject RJW. Plot symbols: *-facilitory bias (bias and step have same sign), o-zero bias, x-inhibitory bias (bias and step have opposite sign). Regression lines labeled with values of bias in kgm.

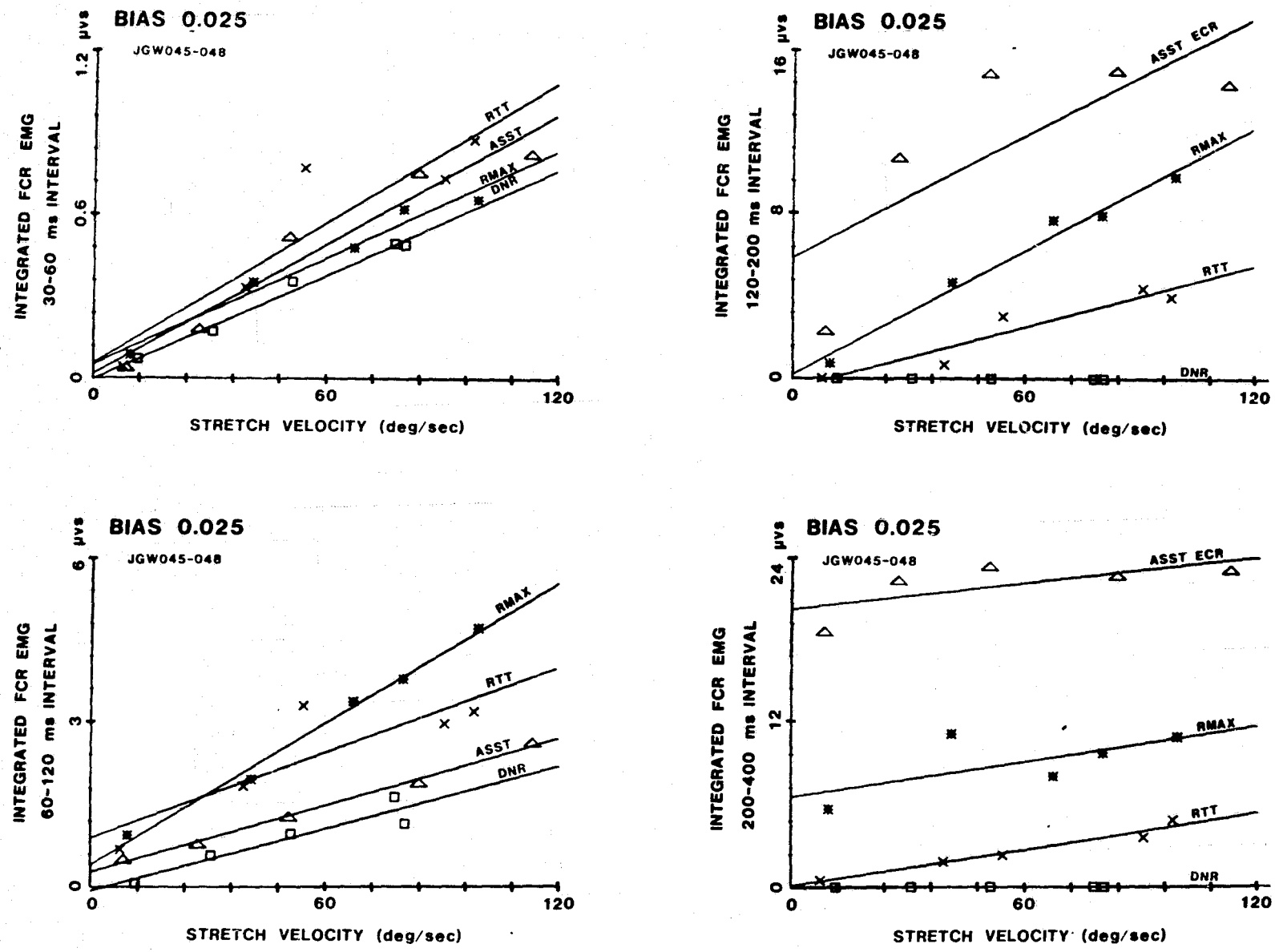


Figure 6: Gain plots for all intervals showing effects of instruction. Data from FCR of subject JGW. Plot symbols: x-RMAX, *-RTT, ^-ASST, o-DNR. Note that ASST data for 12Ø-2ØØ and 2ØØ-4ØØ ms intervals is from ECR.

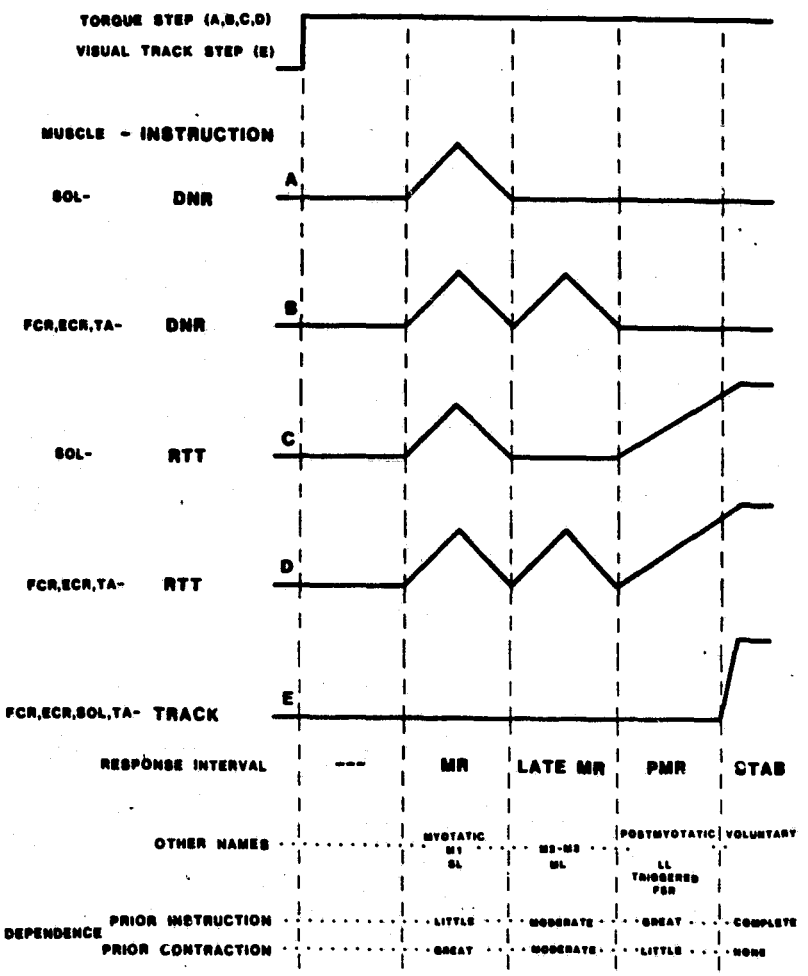


Figure 7: Generalized scheme for the classification of EMG responses to various stimuli. Step trace (top) represents step torque perturbation for EMG responses in traces A, B, C, and D, and visual step tracing stimuli for EMG response in trace E. Abbreviations: RTT-react to target, RMAX-react maximally, DNR- do not react, ASST-assist, SOL-soleus, FCR-flexor carpi radialis, ECR extensor carpi radialis, MR-myotatic response, PMR-postmyotatic response, STAB- stabilizing response. See text for abbreviations of other names. Time scale arbitrary.

MANUAL CONTROL ANALYSIS OF DRUG EFFECTS ON DRIVING PERFORMANCE*

By Alison Smiley, Kenneth Ziedman and Herbert Moskowitz

Southern California Research Institute

SUMMARY

Effects of secobarbital, diazepam, alcohol and marihauna on car-driver transfer functions obtained using a driving simulator were studied. The first three substances, all CNS depressants, reduced gain, crossover frequency and coherence which resulted in poorer tracking performance. Marihuana also impaired tracking performance but the only effect on the transfer function parameters was to reduce coherence.

INTRODUCTION

Manual control analysis of tracking performance has been most frequently used to study changes in task variables, such as controlled element dynamics, rather than operator variables such as fatigue, or drug effects. This paper addresses the issue of drug effects on the tracking behavior of human operators, a topic of great interest because of its importance to traffic safety. The drugs studied were secobarbital, a sedative hypnotic, diazepam, a widely prescribed sedative minor tranquilizer, and marihuana, a common recreational drug. The diazepam and marihuana were combined with alcohol.

The transfer function approach for examining drug effects was chosen for two reasons. First, it allows some differention between drugs in terms of which aspects of perceptual-motor behavior are being affected. Second, it provides a unified framework in which to interpret results, unlike the interpretation of an assortment of tracking parameters.

There have been few published studies of drug effects on transfer function parameters. Allen et al (reference 1) using both an instrumented car and a driving simulator found the effect of alcohol was to reduce the driver's gain and decrease the coherence (i.e., the driver became a less responsive and less linear tracker). Reid (reference 2) calculated describing functions for subjects performing a compensatory tracking task under the influence of alcohol alone and in combination with marijuana. Similar effects of reduced gain and reduced linearity of response were found as well as an increase in subject internal processing time. All these changes in control

* This work supported by The National Institute on Drug Abuse and The National Highway Traffic Safety Administration under NIDA Contract No. 271-76-3316

contributed to degraded tracking performance.

Tracking task measures related to transfer function parameters have been used to distinguish between the effects of different drugs. Smiley et al. (reference 3) compared the effects of alcohol alone and in combination with diazepam, diphenhydramine (an antihistamine) and marihuana on driving in an instrumented car. The power spectra of steering wheel angle, a measure of the amplitude of steering movement over a range of frequencies, showed: (1) that subjects made larger steering movements under the alcohol + diphenhydramine, alcohol + marihuana and alcohol alone treatments when compared to placebo, and smaller movements under the alcohol + diazepam treatment and (2) that subjects' steering movements were slower under all drug treatments, and slowest for the alcohol + diazepam condition, when compared with placebo.

PROCEDURES

Drug Treatments: The results reported in this paper are derived from three separate experiments on the effects of drugs on human performance. The first experiment examined secobarbital alone at three dose levels: 0, 1.1 and 2.2 mg/kg bodyweight using 15 subjects. The second and third experiments were drug-alcohol interaction studies with the drug and the alcohol each tested at three levels. A separate group of 15 subjects was tested at each alcohol level, making a total of 45 subjects tested in each of the two studies. Each subject received all three drug doses. Dose levels used in the second experiment were: diazepam: 0, 0.11, 0.22 mg/kg bodyweight and alcohol: 0, 0.51, and 1.02 gms/kg bodyweight. For the third experiment dose levels were: marihuana: 0, 100, 200 mcg THC/kg bodyweight and alcohol: 0, 0.425 and 0.68 gms/kg bodyweight. The alcohol was administered in a vodka-orange juice mixture, the diazepam and secobarbital by capsules, and the marihuana by smoking.

The dose levels of secobarbital and diazepam were those generally used in therapeutic situtations. The highest dose level of alcohol produced a blood alchol concentration (BAC) of 0.11%, just over the 0.10% BAC legal presumptive limit for impariment in California. A questionnaire, given to the subjects who received marihuana, showed that the 100 and 200 mcg THC/mg bodyweight doses produced the same "high" as the subjects experienced in their social use of marihuana between "less than half the time" and occasionally".

Subjects: Participants met the following criteria: male, 21-45 years old, 61.5-91 kg bodyweight, 20/30 minimum vision in each eye, moderate to light heavy alcohol use as defined by the Cahalan et al (reference 4) scale, and having at least three years driving experience. Subjects were screened using a medical examination and a standardized personality test for possible physical or emotional counter indications.

Testing Schedule: Subjects attended three training days within a

three-week period. On each training day subjects completed two forty-five minute simulator runs. After training, subjects attended three treatment sessions separated by two week intervals. At each treatment session, a subject was given an eight-minute "warm-up" run in the simulator, after which dosing began. Sixty-five minutes after the start of dosing the subject began a 45-minute simulator run. The testing time was chosen so that the drug and alcohol blood levels would peak during the run. Measurements of blood alcohol concentration were taken before dosing to insure an initial 0% BAC, just before and just after testing, and every hour until the BAC was below 0.03%. Pulse rates and blood samples to determine drug levels were also taken during these experiments but the results will be reported elsewhere.

Apparatus: The driving simulator used in this study was developed to test performance of control and decision skills shown to be both critical to the driving task, and sensitive to drug effects. The simulator design was based on a general purpose digital computer (PDP 11/60) and associated graphics system (Megatek 7000) which provided:

* implementation of realistic vehicle dynamics
* generation of a roadway (straight, curved, etc.) and roadway elements (signs, obstacles, other vehicles, etc.) and
* data recording and analyses.

Detailed descriptions of the simulator are given in Michaelson et al. (reference 5) and Allen et al (reference 6). A number of tasks were performed during the run including curve negotiation, passing maneuvres and emergency stops. The results reported in this paper were derived from one task, the wind-gust control task, which was presented three times and lasted approximately 2 minutes each time. In this task the driver was required to keep the simulator centered in the lane at a constant speed of 80 k.p.h. while being buffeted by simulated wind gusts. A heading angle disturbance signal consisting of a sum of seven sine waves was used to create the wind gust effect and allowed the derivation of car-driver transfer functions.

THEORETICAL ANALYSIS

Figure 1 shows the control loop structure which represents the driving simulator. The driver is assumed to use heading angle (ψ) and lateral position (y) inputs to steer the driving simulator. The nested loop structure shown, an inner loop operating on heading angle and an outer loop on lateral position, has been successfully fitted to data from experienced drivers by Weir and McRuer (reference 7). By keeping the inner loop closed the driver can operate on lateral position error with a simple gain, i.e., corrections of lateral position may be facilitated by means of heading angle corrections. The sum of sines disturbance, denoted ψ_d, contained the following seven frequencies: 0.553, 0.916, 1.288, 2.023,

2.947, 4.235 and 5.705 radians per second, i.e., the spacing was at approximately equal intervals over a logarithmic scale.

$G_\delta^V(jw)$ and $G_\delta^\psi(jw)$ represent the vehicles dynamics for lateral velocity and heading angle respectively (jw is the complex frequency variable). In the simulator the vehicle dynamics are simulated digitally by differential equations. These equations represent a mid-size American sedan and were drawn from a study determining vehicle dynamics of various cars by McRuer et al. (reference 8).

Using Figure 1 it may be seen that the system equations can be written as:

$$\psi = Y_\psi \, G_\delta^\psi \, (n - Y_y y - \psi) + \psi_d$$

and

$$y = \frac{U_o}{s} \psi + Y_\psi \frac{1}{s} G_\delta^V \, (n - Y_y y - \psi)$$

where U_o represents the forward speed of the simulator and n, the remnant, that part of the driver's steering input uncorrelated with the heading angle disturbance.

The next steps in the derivation of the car-driver transfer function are:

1) substitute for lateral position, y, in the first equation using the second equation.
2) write the equation with the effective open loop system gain,

$$\frac{\psi - \psi_d}{\psi}$$

on the left side. The heading disturbance ψ_d, is considered the system input and the heading angle ψ, is considered the output.
3) cross-correlate each side of the equation with ψ_d to obtain

$$\frac{{}^\Phi\psi_d\psi - {}^\Phi\psi_d\psi_d}{{}^\Phi\psi_d\psi}$$

The cross-correlation function ${}^\Phi\psi_d\psi(jw)$ describes the general dependence of the heading angle signal ($\psi(t)$) on the heading angle disturbance signal ($\psi_d(t)$) in terms of amplitude and phase relationships. The process of obtaining the cross-correlation function involves converting $\psi(t)$ and $\psi_d(t)$ to the frequency domain using Fast Fourier transforms. The heading angle and heading angle disturbance signals were recorded during the

wind-gust control task at a rate of 7.5 times per second. (Analysis techniques are described in detail by Bendat and Piersol (reference 9).) Cross-correlating the remnant, n, with ψ_d causes all remnant terms to disappear as $^{\Phi}\psi_{d.n}=0$ by definition. After cross-correlation, the effective open loop system gain is represented by the function $\frac{^{\Phi}\psi_{d}\cdot\psi - ^{\Phi}\psi_{d}\psi}{^{\Phi}\psi_{c}\psi}$.

By definition the open loop gain equals the product of the driver transfer function Y_p and the car (simulator) transfer function Y_c;

that is, $Y_pY_c = \frac{^{\Phi}\psi_{d}\psi - ^{\Phi}\psi_{d}\psi}{^{\Phi}\psi_{d}\psi}$.

RESULTS

Obtained blood alcohol concentrations were 0.06% and 0.11% in the diazepam-alcohol interaction study, and 0.05% and 0.08% in the marihuana-alcohol interaction study.

Car-driver transfer functions were calculated and averaged for each group of 15 subjects for each drug, and for each alcohol level tested (in some cases fewer subjects were available). For all drug and alcohol treatments there were significant increases in tracking error ($p<0.05$). Considering the results for each of the three drugs under the no alcohol, or placebo alcohol condition, the largest increase in tracking error was found for the high dose of secobarbital, the sedative hypnotic (see table 1). The increases in tracking error produced by the high doses of diazepam (sedative tranquilizer) and marihuana were approximately equivalent, and half that found for the secobarbital high dose.

The tracking error results for the various alcohol doses are not as clear because different groups of subjects are being compared. Also initial differences between groups, exacerbated by running the active alcohol groups some months after the placebo alcohol group had been completed, make the alcohol results from the marihuana-alcohol interaction study less reliable than they might be. In the diazepam-alcohol interaction study tracking performance under the 0.06% BAC-placebo drug condition was about the same as under the high dose diazepam-placebo alcohol condition. Tracking error appeared to be linearly related to alcohol dose, and doubled for the 0.11% BAC condition in comparison with the 0.06% BAC condition.

Figure 2 shows average car-driver transfer functions obtained for the three dose levels of secobarbital. There were large drops in gain and phase angle with increasing dose. Gain at all frequency points, and crossover

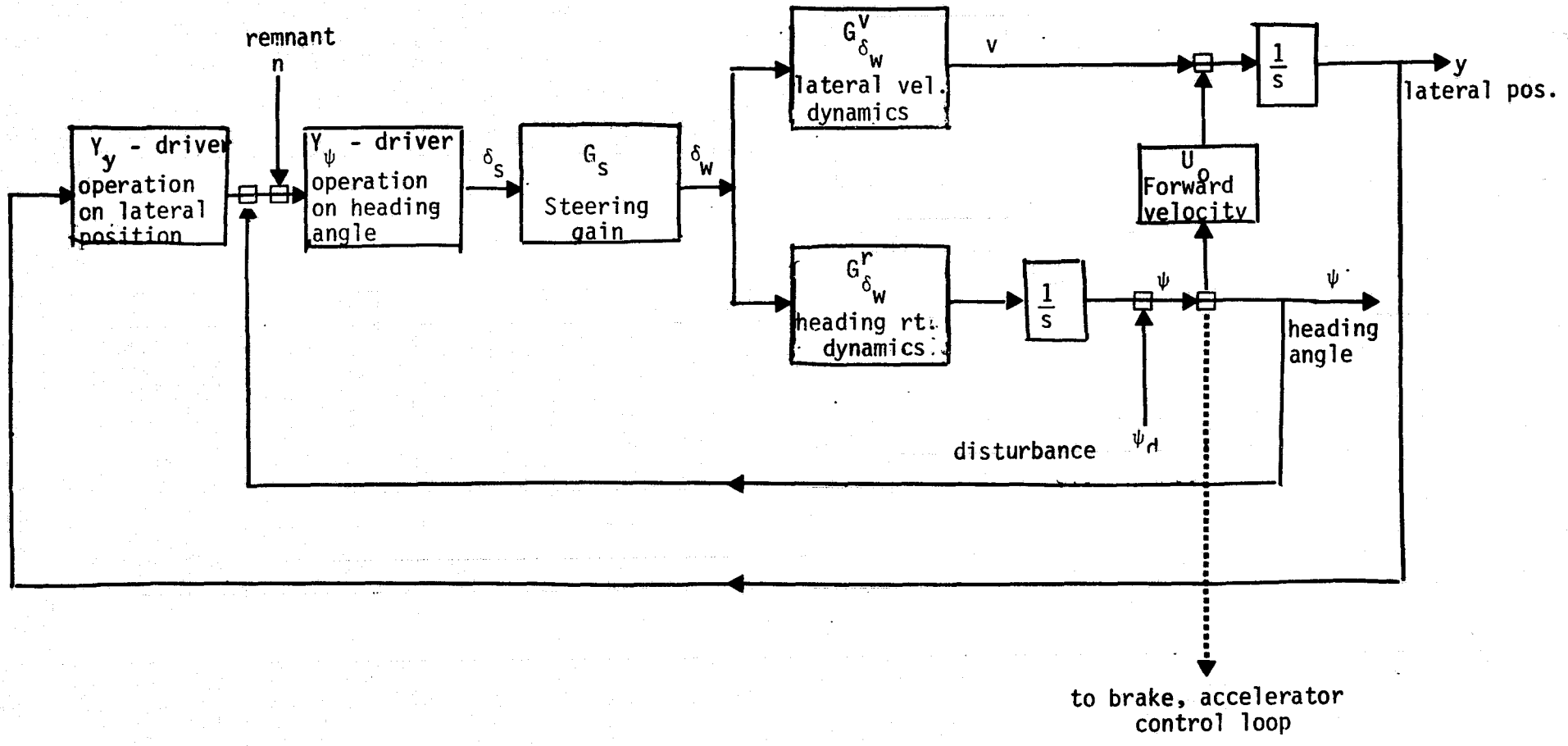


FIGURE 1: Control loop diagram for driver/vehicle system

frequency, were significantly reduced ($p<0.05$). Coherence between the heading angle disturbance and the heading angle signal was significantly reduced ($p<0.001$). Phase margin, which is an indicator of system stability, was unaffected (see Table 1).

The diazepam results for the placebo alcohol group are shown in Figure 2. Again there were significant drops ($p<0.05$) in gain at all frequency points and in crossover frequency. However, the average drop in gain was less for diazepam than for secobarbital treatment. Coherence was also significantly reduced by the diazepam treatment.

Alcohol results for placebo diazepam condition only are shown in Figure 3. These results are not as quite clear cut as for the other drugs. Here the lowest gains were for the 0.06% BAC alcohol condition. The effect of alcohol was significant ($p<0.05$) on the gains at the second, third and fourth frequency points only, with the Newman-Kuhls comparison of means test showing the placebo alcohol gains to be significantly higher than for either the 0.06% or 0.11% BAC conditions. Coherence was significantly reduced by the alcohol treatment, the greatest drop being for the 0.11% BAC condition. Mean crossover frequency also was significantly reduced ($p<0.05$). (It should be noted that the alcohol comparison was between different groups of 15 subjects while the drug comparison were based on repeated measures with 45 subjects. Thus the test of the alcohol effect was not as strong as the test of the drug effect.)

In summary, the results from the secobarbital, diazepam and alcohol treatments are much the same. Gains were significantly reduced as were crossover frequency and coherence. Phase margin was unaffected.

The effects on the car-driver transfer function for the marihuana treatment were very different from the other drugs (see Figure 3). There were no noticeable effects on gain, crossover frequency or phase margin. The main effect appeared to be on coherence which was significantly reduced ($p<0.05$), at the high dose level only. Tracking error was also significantly affected at the high dose only. Despite the fact that tracking error increased the same amount for both diazepam and marihuana, the effects on the car-driver transfer function were very different.

DISCUSSION AND CONCLUSION

The alcohol effects found in this experiment are supported by Allen et al (reference 1) and Reid and Ibrahim (reference 2) who also showed alcohol to reduce gains and coherence in driver transfer functions. Reid and Ibrahim found marihuana to have little effect on amplitude or phase margin but to decrease coherence, similar to the results obtained in this study. No comparative data are available for the secobarbital or diazepam treatments.

TABLE 1

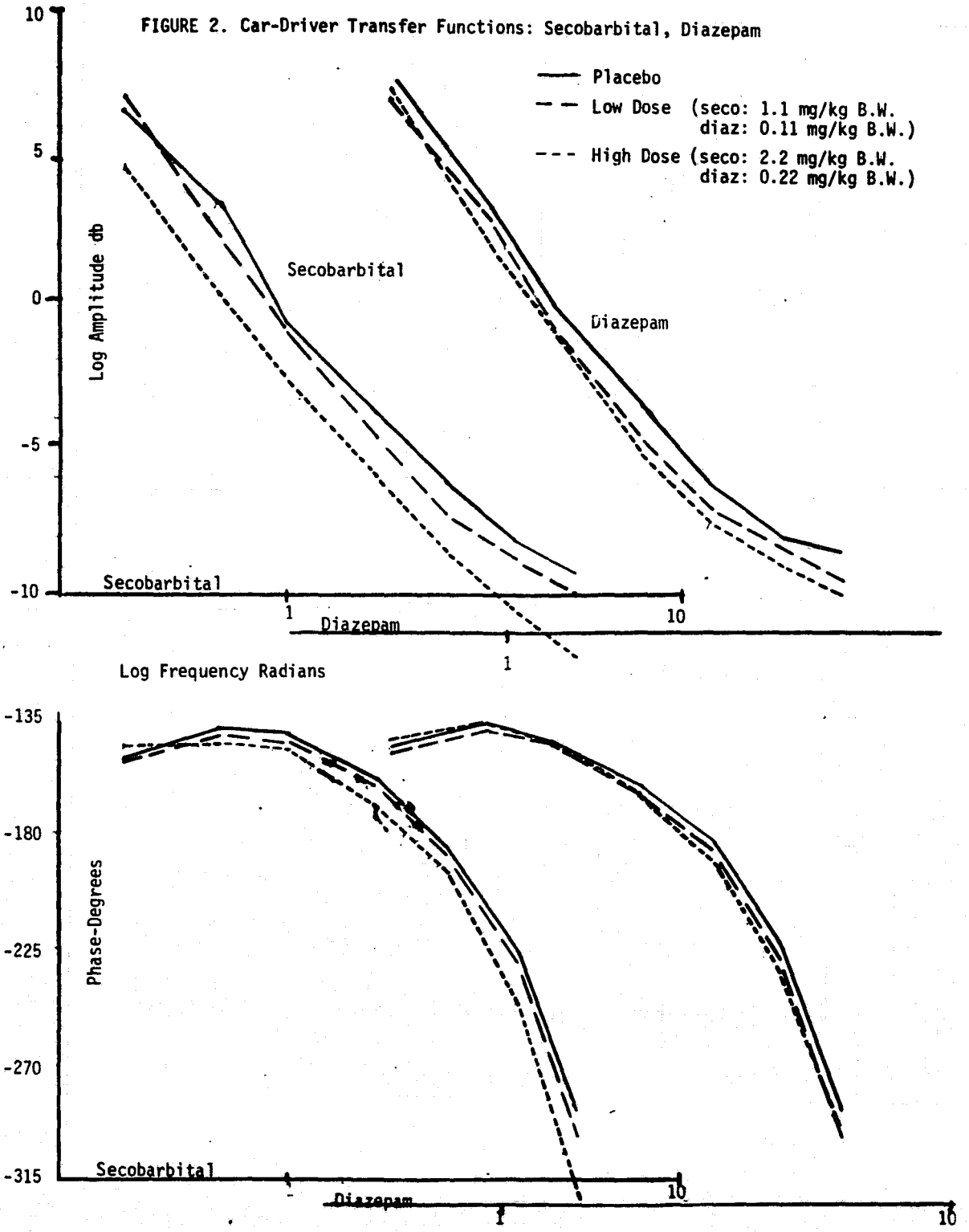
Summary of Means and Analysis of Variance Results

	Secobarbital		Diazepam - Alcohol					Marihuana - Alcohol				
		p_2 value	BAC	0%	0.06%	0.11%	p value diaz alc	BAC	0%	0.06%	0.11%	p value mar alc
Standard Deviation Lateral Position	P_1 1.48	***	P	1.36	1.60	1.86	** ***	P	1.41	1.36	1.70	** —
	L 1.59		L	1.42	1.66	2.17		L	1.42	1.55	1.59	
	H 1.96		H	1.55	1.92	2.16		H	1.68	1.47	1.69	
Coherence	P 0.900	***	P	0.913	0.869	0.829	* ***	P	0.907	0.901	0.869	* —
	L 0.905		L	0.904	0.874	0.808		L	0.909	0.884	0.901	
	H 0.846		H	0.876	0.826	0.821		H	0.863	0.879	0.883	
Crossover Frequency	P 0.216	***	P	0.212	0.168	0.193	*** *	P	0.219	0.188	0.168	— **
	L 0.200		L	0.916	0.156	0.142		L	0.208	0.193	0.171	
	H 0.170		H	0.189	0.167	0.146		H	0.863	0.879	0.883	
Phase Margin	P 0.580	—	P	0.594	0.696	0.593	— —	P	0.636	0.696	0.608	— —
	L 0.597		L	0.601	0.640	0.591		L	0.663	0.656	0.699	
	H 0.577		H	0.587	0.571	0.637		H	0.634	0.682	0.655	

Drug Dose Levels: 1. P = placebo, L = low dose (seco: 1.1mg/kg B.W.; diaz:0.11 mg/kg B.W ;mar 100mcg/kg B.W.
H = high dose (seco: 2.2mg/kg B.W.; diaz:0.22mg/kg B.W.;mar: 200 mcg/kg B.W.

2. p values * $p<0.05$, ** $p<0.01$, *** $p<0.001$

FIGURE 2. Car-Driver Transfer Functions: Secobarbital, Diazepam



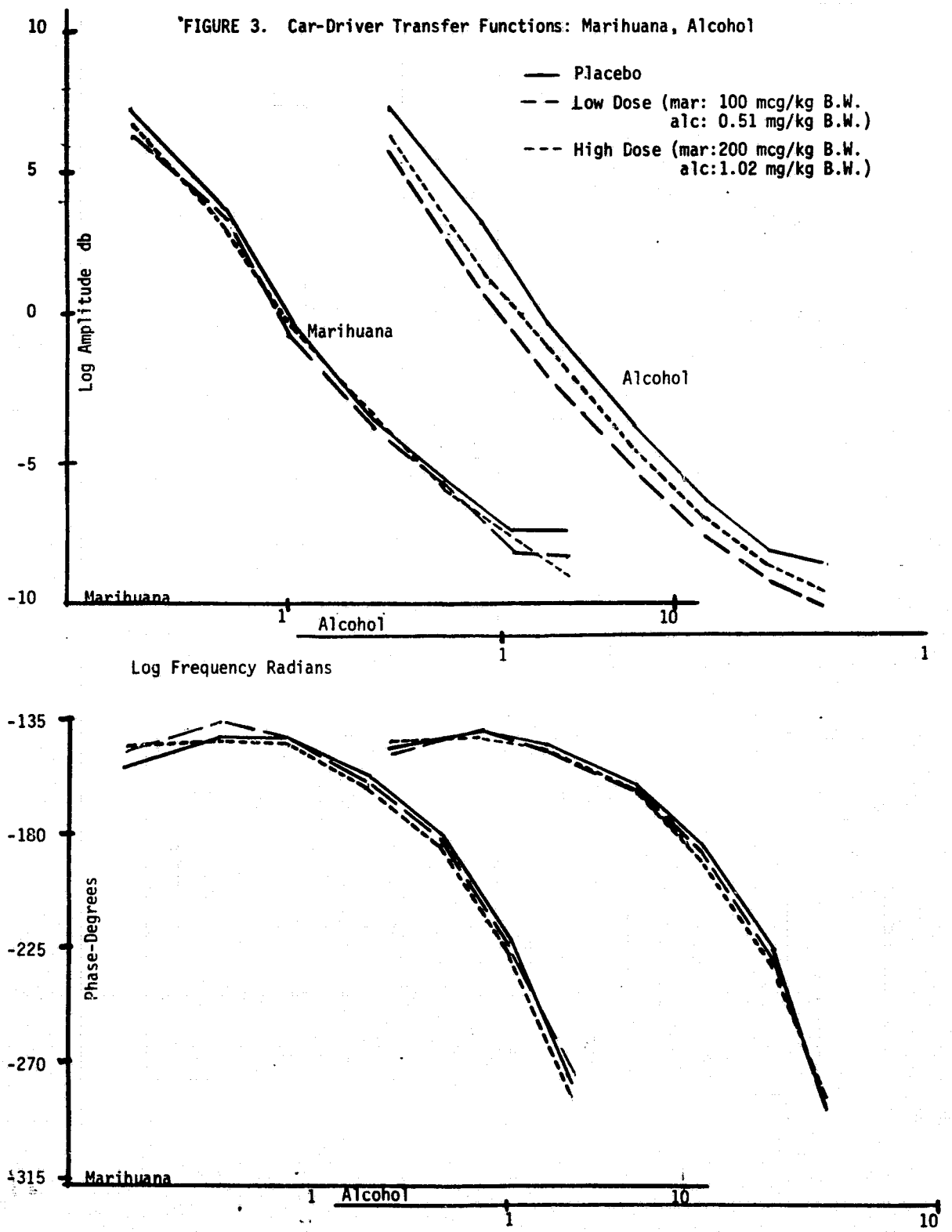


FIGURE 3. Car-Driver Transfer Functions: Marihuana, Alcohol
Placebo
Low Dose (mar: 100 mcg/kg B.W.
alc: 0.51 mg/kg B.W.)
High Dose (mar:200 mcg/kg B.W.
alc:1.02 mg/kg B.W.)
10
5
0
-5
-10
Log Amplitude db
Marihuana
Alcohol
Marihuana
1
10
Alcohol
1
1
Log Frequency Radians
-135
-180
-225
-270
-315
Phase-Degrees
Marihuana
1
Alcohol
10
1
10

Both secobarbital and diazepam act as sedatives, however, secobarbital is generally prescribed at a high enough dose level that it acts as a hypnotic (induces sleep) while diazepam's anti-anxiety properties become evident when it is prescribed at a lesser dosage. Both drugs are CNS depressants as is alcohol. In contrast to these three drugs, marihuana is classified as a psychedelic, i.e., affecting the thought processes. The transfer function analysis clearly distinguished between these two classes of drugs. The analysis also discriminated the degree of sedative effect in that gains were reduced much more for the secobarbital treatment than the diazepam treatment at the dose levels used in this experiment. In addition it is interesting to note that the one drug of the four which is supposed to most affect thought processes is the one whose only transfer function effect was an increase in non linear behavior.

REFERENCES

1. Allen, R.W.; Schwartz, S.; Stein, A.; Magdelano, R. and Hogge, J.: The Effects of Alcohol and Marihuana on Driver Control Behavior. Volume 1 Laboratory Simulation Experiment. Report submitted to National Highway Traffic Safety Administration, under contract No. DOT-HS-501257, April, 1978.

2. Reid, D.; Ibrahim. M.K.F.: The Application of Human Operator Describing Functions to Studies on the Effects of Alcohol and Marihuana on Human Performance. IEEE Transactions on Systems, Man and Cybernetics, Vol.SMC-5, No.5, September, 1975.

3. Smiley, A.; LeBlanc, A.E.; French, I.W.; and Burford, R.: The Combined Effects of Alcohol and Common Psychoactive Drugs: Field Studies with an Instrumented Automobile. Canadian Society of Forensic Science Journal, Vol.8., No.2., 1975.

4. Cahalan, D.; Cisin, I.H.; and Crossley, H.M.: American Drinking Practices, A National Study of Drinking Behavior and Attitudes. College and University Press, 1969.

5. Michelson, S.; Niemann, R.; Olch R.; Smiley, A.M; and Ziedman, K.: A Driving Simulator for Human Performance Studies Using All Digital Techniques. Southern California Research Institute Technical Report, January, 1979.

6. Allen, R.W.; Klein, R.H.; and Ziedman, K.: Automobile Research Simulators: A Review and New Approach. Transportation Research Record 706, Simulation Technology and Traffic Accident Records System, Transportation Research Board, National Academy of Science, Washington, D.C., 1979.

7. Weir, D.H.; and McRuer, D.T.: Measurement and Interpretation of Driver Steering Behavior and Performance. SAE Paper No.730098, 1973.

8. McRuer, D.T.; and Klein R.H.: Automobile Controllability-Driver/ Vehicle Response for Steering Control. Sytems Techonology, Inc. Los Angeles, Submitted to National Highway Traffic Safety Administration under contract No.DOT-HS-359-3-762, November, 1974.

9. Bendat, J.S.; and Piersol, A.G.: Random Data: Analysis and Measurement Procedures. Wiley-Interscience, 1971.

DEVELOPMENT OF A "STANDARD" EMG MEASUREMENT SYSTEM

By D. Antonelli, D. Hary,
and G.A. Bekey

Rancho Los Amigos Hospital
and University of Southern California

ABSTRACT

Considerable controversy exists in the literature concerning the relation between EMG signals and the force generated by a skeletal muscle. Some of the controversy relates to the lack of standardization in the instrumentation and in EMG processing. In an attempt to help solve this problem, a study of EMG electrodes and amplifiers was undertaken. Mathematical models of both wire and surface electrodes were constructed from measurements with a range of frequencies from 50 Hz to 2000 Hz. These models were combined with a mathematical model of a high performance differential amplifier. An ECAP analysis was then undertaken to determine the sensitivity of the amplifier common mode rejection ratio and gain to various electrode configurations and amplifier parameters. The study provides guidelines for the minimum Z_{in} which yields satisfactory performance for a given set of electrodes.

Typical results for the surface electrodes are given in Fig. 1 which shows the gain vs frequency for different input parameter values. As the amplifier's input impedance was reduced, its power gain at low frequencies was dramatically reduced as well. This reduction in gain significantly altered the overall transfer function of the system for all types and sizes of electrodes. In fact, the low half power frequency shifted considerably to the right from its theoretical 0 value to 70 Hz. Since most of the power obtained from surface electrodes is concentrated at frequencies below 200 Hz, a frequency band which does not include a considerable part of this power range would alter the detected power spectrum. In other words, low common and differential mode input resistances in the amplifier will cause erroneous estimates of the EMG spectrum at low frequencies. Furthermore, the decrease in gain with decreasing frequency for this amplifier configuration has the characteristics of a differentiator, which will alter the shape of the input signal in the time domain. A similar result was presented by Geddes (1972) for electrocardiograph amplifiers. Similar results apply to wire electrodes.

The common mode rejection capacity of the system was also degraded by low amplifier input impedance values when the source was unbalanced. This effect was also most pronounced at low frequencies, where the reduction in the CMRR was 40 db. Such a reduction would give rise to increased line frequency noise and hence distort the shape of the measured EMG spectrum. The input guard effect increased the CMRR by 20 db throughout the fre-

quency range for the high CM impedance amplifier. However, with the low CM impedance amplifier, the effect was negligible at low frequencies and had a value of only 4 db at 2000 Hz.

We believe that the results of this study will form the basis for the specification of a standard for EMG amplifiers.

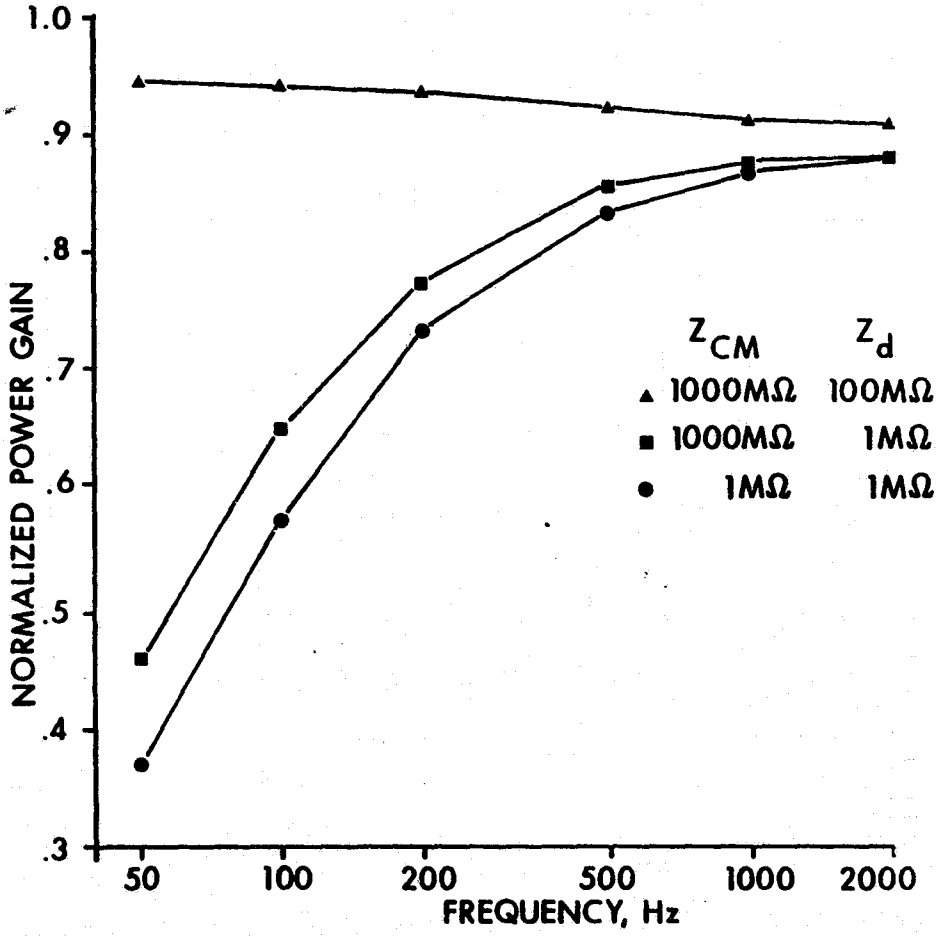


Fig. 1

D45

Pilot Scanning Patterns while Viewing Cockpit Displays of Traffic Information

By Stephen R. Ellis and Lawrence Stark

NASA Ames Research Center &
U.C. Berkeley Department of Physiological Optics

ABSTRACT

Scanning eye movements of airline pilots were recorded while they judged air traffic situations displayed on cockpit displays of traffic information (CDTI). The observed 1st order transition patterns between points of interest on the display showed reliable deviations from those patterns predicted by the assumption of statistical independence. However, both patterns of transitions correlated quite well with each other. Accordingly, the assumption of independence provided a surprisingly good model of the results. Never the less, the deviation between the observed patterns of transition and that based on the assumption of independence was for all subjects in the direction of increased determinism. Thus, the results provide objective evidence consistent with the existence of "scanpaths" in the data.

INTRODUCTION

In the following experiment we examine the spatio-temporal structure of scanning eye movements made by airline pilots while viewing a cockpit display of traffic information (CDTI) previously studied by Palmer, Jago, et. al. (2, see also ref. 6). While viewing the displays, the pilots had to determine if intruding aircraft would pass in front or behind their ownship. In particular we wished to examine the specific sequencing of fixations made with the goal of determining an information processing model of the in front/behind decisions the pilots made. Other aspects of the scanning will be discussed in future more detailed papers(see ref 5).

METHODS

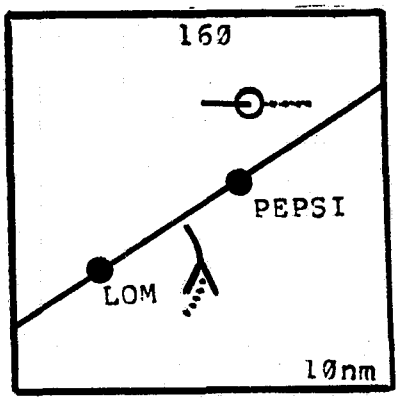
Display conditions

A series of .24 track-up CDTI displays was generated on a calligraphic computer graphics system in the manner of Palmer et. al. (2,3). Each display depicted an encounter between two aircraft at the same altitude. (see inset of figure 2). The miss distances for all encounters were set at 6000 feet while

ownship position was updated every 0.1 second and and the intruder was updated every 4 seconds. Map range was set at 18.5 km (10 nm). Each encounter consisted of 7 4-second updates; 2 before the intruder appeared, and 5 afterwards. The display was blanked in such a way that after the last update, there remained 44 seconds before the fight paths of the aircraft crossed. After the display was blanked, the subject was prompted to decide if the intruder would pass in front or behind ownship.

The encounters represented both straight and turning horizontal encounter geometries. Turn rate was constant at 1.5 degrees/second. All aircraft had 32 seconds of previously tracked positions displayed as 8 dots of trail, one for each update, and had a 32 second predictor (see insert of figure 1 for sample display)

Figure 1



Two sets of identical encounters were prepared: in one all aircraft had "straight" predictors which were based on extrapolation of current ground speed, in the other "curved" predictors were used which were based on current ground speed and turn rate. In addition to ownship and the intruder the display contained two geographical locations (LOM and PEPSI) and a route shown as a solid line. All trajectories crossed near LOM, a fact neither pointed out to the subjects nor discovered by them during the course of the experiment.

Direction of gaze data were recorded with a Gulf and Western 1994 pupilometer-based eye monitor which was calibrated by recording fixations at 25 reference points in a 5 X 5 array, 14 degrees/side, centered in the subjects forward field of view. The display itself subtended a rectangle 12 by 10 degrees at the viewing distance of 75 cm. The eye monitor output (x,y direction of gaze and pupil diameter), the subjects signals, and the time markers from the videotape were all digitally recorded. The sample rate was 33 hz, the maximum allowable by the eye monitor.

Subjects

Eight airline pilots were subjects in the experiment. All had had at least 3 hours experience in previous experiments using similar CDTI displays and requiring similar in front behind judgements. All had at least average performance in these previous studies.

Procedure

During an orientation session before each experiment, the subjects were told that the purpose of the experiment was to determine if pupilary changes could be used to predict their in front/behind judgements. Lengthly briefing was unnecessary due to their extensive familiarity with the CDTI project in general and the display format in particular; the meaning of all parts of the symbology was, however, reviewed and each subject was given about 20 minutes practice making in front/behind judgements before their scanning patterns were recorded. After the initial practice, the eye monitor was adjusted to track the left eye An initial 25 point calibration was taken by having the subject signal when he had fixated each of the reference points. The median x,y position taken during each of these fixations provided the basis for subsequent linearization and removal of cross-talk. Interspersed between data gathered during the encounters were reset calibrations taken to correct for drift by having the subject refixate a position corresponding to the center of the calibration grid. A reset calibration was taken whenever the signal was observed to drift more than about 1.0 degree.

Data Processing

After the experiment, the data were transferred to a PDP-11/70 computer where it was linearized according to a piecewise-linear approximation derived from the calibration fixations. Fixation locations were determined by passing a space-time window through the data and identifying fixations as those clusters of temporally contiguous datapoints representing at least 90 msec (3 samples) falling within a square region 0.82 degrees(side, approximately the resolution of the eye monitor. If a cluster of data points potentially considered a fixation were interrupted by less than 90 msec of saturated values, they were considered as a single fixation. This protected fixations from being interrupted by blinks. Longer interruptions resulted in their being considered as two separate fixations.

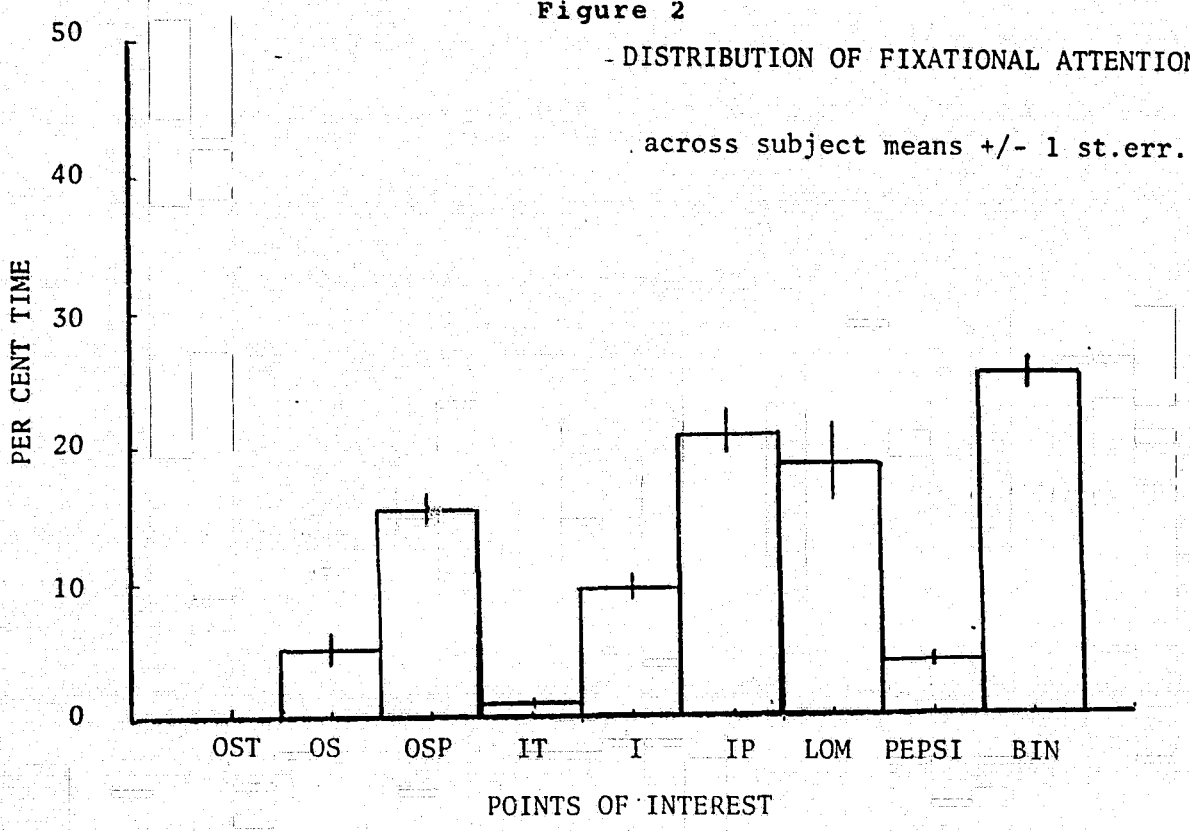
After identifying the positions, duration, and onset time of all fixations, the data were correlated with records of the positions of all points of interest as a function of time after the beginning of each encounter. Thus, each fixation could be assigned to one of eight possible points of interest: the end of ownship's trail (OST), ownship present position (OS), the

end of ownship's predictor (OSP), the end of intruder's trail (IT), intruder's current position (I), the end of the intruder's predictor (IP), location PEPSI (PEP), and location LOM (LOM). (see figures 1 and 2) All fixations not within 2 degrees of any of the above points of interest were assigned to the BIN. The data were then tabulated to determine overall distribution of fixation duration as well as separate distributions for each point of interest. Per cent of time spent at each point of interest was determined.

RESULTS

The pilots distributed their fixational attention across the the eight points of interest in a highly stereotyped manner as indicated in figure 2.

Figure 2



Their distributions of attention reflect the differential usefulness of the spatial information at each point of interest for the in front/behind judgement that they were required to make. Earlier studies of this particular display had shown, for example, that the presense or absence of the trail had no effect on accuracy of the pilots judgements, whereas the presense and type of predictor was very important (2). The higher proportion of viewing time on the location LOM probably occured because it was the point of intersection of the flight paths for all the encounters.

The differential probability of viewing the various points of interest leads to a constraint on the scanning sequences that could produce the impression of sequential scanning, namely that transitions between high probability points of interest are likely simply due to the zero order probability of viewing the respective points. Accordingly, any claim for the observation of repetitive sequences in the the transition patterns among points of interest must first show that the extent of the "sequenciness" exceeds that which could be produced by the zero order probabilities.

We have examined this possibility by using a method described by Senders et. al, (4). They note that the joint assumption of 1) statistical independence of the transitions and 2) the existence of unobserved transitions from each point of interest to itself provides a means of calculating p(a to b), probability of a transition between any two points of interest, provided p(a), p(b), the zero order probabilities of the two points are known.

$$p(a \text{ to } b) = \frac{p(a)p(b)}{1 - \sum_{i}^{n} p(i)^2}$$

We have used these calculated probabilities to determine expected frequencies of 1st order transitions between points of interest to compare with observed transition frequencies. This comparison was made on a subject by subject basis with a chi-square goodness of fit test on the entire distribution of observed transitions with that of expected transitions. The very low probability of viewing the end of the aircraft's trails, resulted, however, in very low expected frequencies for some of the transitions. Thus, it was necessary to collapse some of the transitions with expected frequencies less than 4.0 in order to insure that not more than 20% of the terms in the chi-square calculation had expected frequencies less than 5.

Table 1

Subj	X-sqr(df)	Number of transitions	Corr	Corr(df) log	H(matrix) observed bits	H(matrix) expected bits
1	35.29(12)p<.01	154	.965	.844(39)	1.606	1.785
2	152.55(21)p<.001	417	.962	.785(51)	1.940	1.980
3	134.78(27) " "	409	.943	.719(43)	1.846	2.115
4	97.28(19) " "	348	.955	.747(49)	1.838	1.978
5	85.80(25) " "	270	.935	.794(43)	2.003	2.197
6	84.54(30) " "	431	.970	.843(57)	2.282	2.509
7	178.81(24) " "	429	.962	.821(48)	1.885	2.104
8	78.26(22) " "	275	.945	.721(43)	2.002	2.183

The chi-square test showed for every subject a highly reliable difference between the observed and expected transition frequencies. The reliablility of this difference was partly due

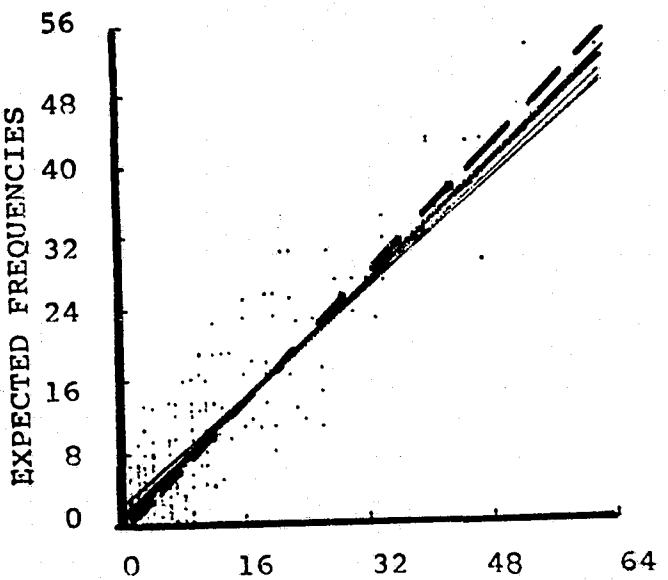
to the very great statistical power of the test since the total number of transitions analyzed for each subject ranged between 154 and 431. Furthermore, the test neither addresses the magnitude nor the direction of the deviation.

Accordingly, in order to assess the extent of the deviation each subject's expected transition frequencies were ploted against his corresponding observed frequencies. In a such a plot a perfect prediction corresponds to a linear regression with a slope of 1.0 and correlation of 1.0. As is clear from figure 3, the slope of the regressions for each subject are quite close to 1.0 (dashed line) and there is a strong linear relation between the observed and expected frequencies.

Figure 3

FREQUENCIES OF TRANSITION 0TH ORDER MODEL

(ALL SUBJECTS)



Two correlations between observed and expected frequencies are shown in table 1 for each subject. The first is the pearson correlation corresponding to the regresssions shown in figure 3. The second is a pearson correlation based on log transforms of both expected and observed frequencies which corrects for the skew in the marginal distributions of observed and expected frequencies.

Though there exists a reliable deviation between the observed and expected frequencies, the independence model provides a surprisingly good prediction of the actual 1st order transition pattern. However, the chi-square test comparing observed and expected transition frequencies provides no indication of the direction of the deviation. The observed distribution could, for example, be more rectangular than that predicted, i.e. the transition frequencies in any row of the 1st order matrix are more equal to each other. Such a deviation would be in the direction of less determinism. Conversely, the observed distribution could be less even with higher peaks and lower valleys; such a deviation would be in the direction of more determinism. The most deterministic case being that with only one type of transition occuring on each row of the 1st order matrix.

We assessed the direction of the actual deviation on a subject by subject basis by treating each ith row of the 1st order transition matrix as a probability vector and calculating H(i), the "information" in bits contained in it. The total "information" in the matrix, H(matrix), was calculated as a weighted average of the rows where n(i), the total number of transitions in each row, providing the relative weights and p(i,j) was the probability of a transition to point j given a previous point of interest i.

$$H(\text{matrix}) = \frac{n(i)\ \left[-\sum_{j}^{n} p(i,j)\log_2 p(i,j) \right]}{N}, \quad N = \sum_{i}^{n} n(i).$$

This measure of uncertainty when applied to each subject's observed and expected transition matrices consistently indicated that the observed transition matrices were more deterministic than the expected matrices (two-tailed sign test $p < .008$). thus establishing the direction of the statistically significant chi-square previously discussed.

DISCUSSION

The above results show that a surprisingly unstructured model of the transition pattern of fixational eye movements can provide an approximation of the actual pattern of transition among points of interest on a simple display. To the extent this approximation is correct the results are consistent with Senders visual sampling model. It predicts the number of looks at a point of interest as a function of the bandpass of the signal presented there. (4). The collection of the eye movement data over an extended period of time and across different display conditions, however, raises the possibility that the data reflect the mixing of a variety of scanning strategies and that the transition pattern of the subjects was statistically nonstationary. This procedure would bias the results against

evidence of deterministic 1st order transition patterns. Thus, the observation that all the subjects transition patterns deviated from that calculated from the independence model in the direction of more determinism is particularly significant and provides some support for the hypothesis of Noton and Stark (1) that visual scanning is characterized by nonrandom repetitive sequences of fixations which they called scanpaths.

Clearly, the way to more explicitly demonstrate these patterns is to study situations in which the covert switching of information processing strategies can be externalized and used to separate the eye movement data corresponding to different strategies. Indeed, the extent of the deviation of the scanning pattern from that expected by statistical independence may reflect different information loads on the pilot. Simple monitoring may be well modeled by a independence model as used above, while more difficult procedure-following may produce more deterministic scanpaths.

REFERENCES

1. Noton, D. & Stark, L.: Scanpaths in saccadic eye movements while viewing and recognizing patterns. Vision Research, 1971, 11, 929-942.

2. Palmer,E.A., Jago,S., Baty,D.L., and O'Connor,S.: Perception of Horizontal Aircraft Separation On a Cockpit Display of Traffic Information, Human Factors, Oct. 1980, 22, 605-620.

3. Palmer, E.A., Jago, S.J. and DuBord, M.: Horizontal Conflict Resolution Maneuvers with a Cockpit Display of Traffic Information. to be presented at the 17th Annual Conference on Manual Control, UCLA, Los Angles, Calif., June 1981.

4. Senders, John W., Grignette, M.C., & Smallwood, R.: An investigation of the visual sampling behavior of human observers. NASA CR 434, January 1966.

5. Stark, L., Ellis, Stephen R.: Scanpaths revisited. Chapter in Eye Movements Perception and Cognition. edited by Dennis Fisher, Richard Monty, and John Senders, Lawrence Earlbaum Associates, Hillsdale, New Jersey, 1981.

6. Verstynen, Harry A. Potential roles for the cockpit traffic display in the evolving ATC system. Proceedings of the International Air Transportation Meeting, Cincinnati, Ohio , May 20-22, 1980, SAE Technical Paper #800736.

PUPILLOMETRY, A BIOENGINEERING OVERVIEW

G. Myers, J. Anchetta, B. Hannaford, P. Peng, K. Sherman, L. Stark, F. Sun, S. Usui, UCB, Berkeley, CA.

INTRODUCTION

We have chosen to study the pupillary control system for two classes of reasons: First there are reasons of convenience. The pupil is exposed to view and is relatively easily measured. Most of it's inputs (light, accommodation, and vergence) are easily controlled.

The second class of advantages is essentially analytic. The inputs are relatively well understood. Light and accommodation/vergence level form the most important inputs to the system. One can study its open loop response to light using a very simple technique-Maxwellian view (Figure 1). Inputs from cortical structures, while present, are less significant in the pupil system than in other biocontrol systems. Thus the pupil functions very nearly as a true reflex. The basic functions of the pupil are to: 1. Control the amount of light on the retina, and 2. Assist the accommodative system by changing the depth of field as required. The pupil has only one degree of freedom, it's size (area or diameter), thus simplifying the state equations.

The above factors mean that the pupil system may be studied without altering its structure. We believe that it is imperative to study the behaviour of a complex system in an intact state. Our studies are conducted on cooperative human subjects.

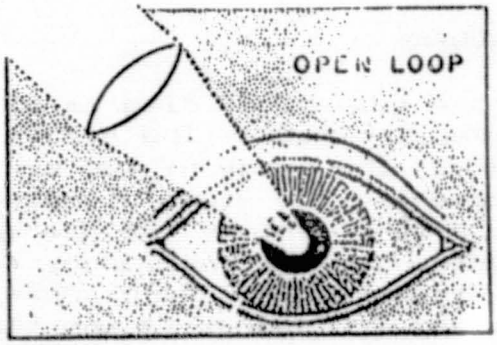


Figure 1. In Maxwellian view, light is focused at the plane of the pupil and passes through the center. Changes in pupil size do not change the amount of light on the retina, thus the pupil control system is open loop.

INSTRUMENTATION

A HISTORY OF THE PUPILLOMETER

The pupil has been studied since 1619.But in that early period the method was direct observation for clinical purposes. In 1927, Lowenstein (Ref 1) used ultra-violet movies and in 1942 used infrared movies (Fig 2). He made motion pictures of the eye and measured the size of the pupil by hand. He obtained accurate data and ploted pupil response curves. But this method required much time and a large amount of infrared film.His interest was also primarily clinical.

The first dynamic real time pupillometer(infrared photo-electronic method, see Fig. 3)was built by Stark in 1957 (2). Using infrared light illuminating the eye surface, the total amount of reflected light depends upon the pupil size .The amount of light can be measured by an photocell. After amplification, the signal of pupil size was recorded immediatly on a chart recorder,and a camera was used for calibration. Using this method Stark measured the transfer function of the pupillary system and set up a third order differential equation which is the first mathamatical model of the pupil.

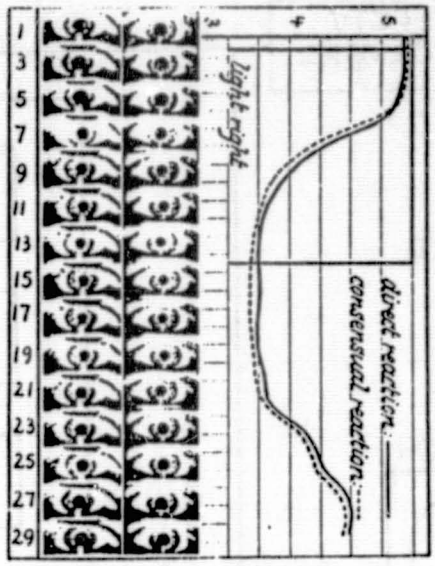
Lowenstein used infrared light passing through a rotating drum, scanning the eye suface, to obtain pupil diameter. Because mechanical scanning is slow, its bandwidth is very limited.

The TV pupillometer was first used by Stark in 1961 to measure the pupil diameter. And now a new computer TV pupillometer measuring the area of pupil has been developed in our lab.

SYSTEM DESCRIPTION

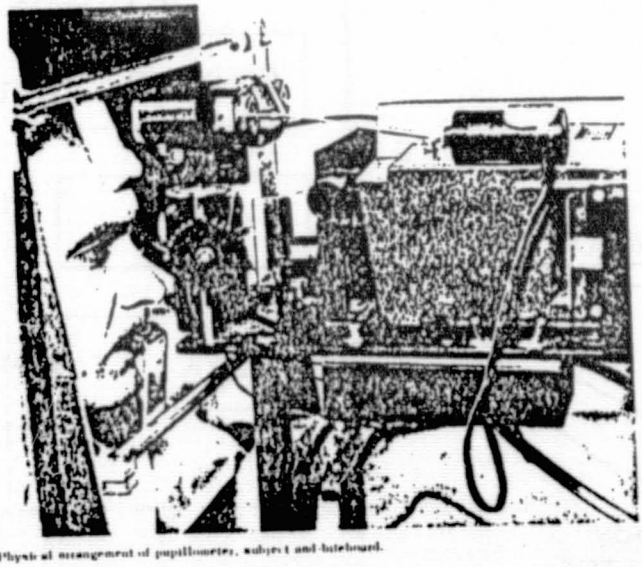
HARDWARE

Figure 4 is a block diagram of the Microprocessor Based Integrative Pupillometer. The eye is illuminated by an infrared source ,the image of the eye is captured by a closed circuit TV camera . The video signal is DC restored , and threshold at a level set by the operator. Generally, the pupil is the darkest object in the image, and no additional processing is required. If dark areas due to eyelashses, eyelids, limbus, etc appear in the thresholded image, they may be excluded by spatial windows,which are also controlled by the operator. The video scans the eye, refreshing the image 60 times/sec. As the video scans the pupil image, a digital counter is enabled. The counter is incremented by a free running clock (18MHz). At the end of a frame,the computer program reads the contents of the counter, and resets the



Typical Photographic Film

FIG. 2 PHOTOGRAPHIC FILM PUPILLOMETER (1942, O. Lowenstein)



Physical arrangement of pupillometer, subject and lateboard.

FIG. 3 DYNAMIC PHOTOELECTRONIC PUPILLOMETER (1957, L. Stark)

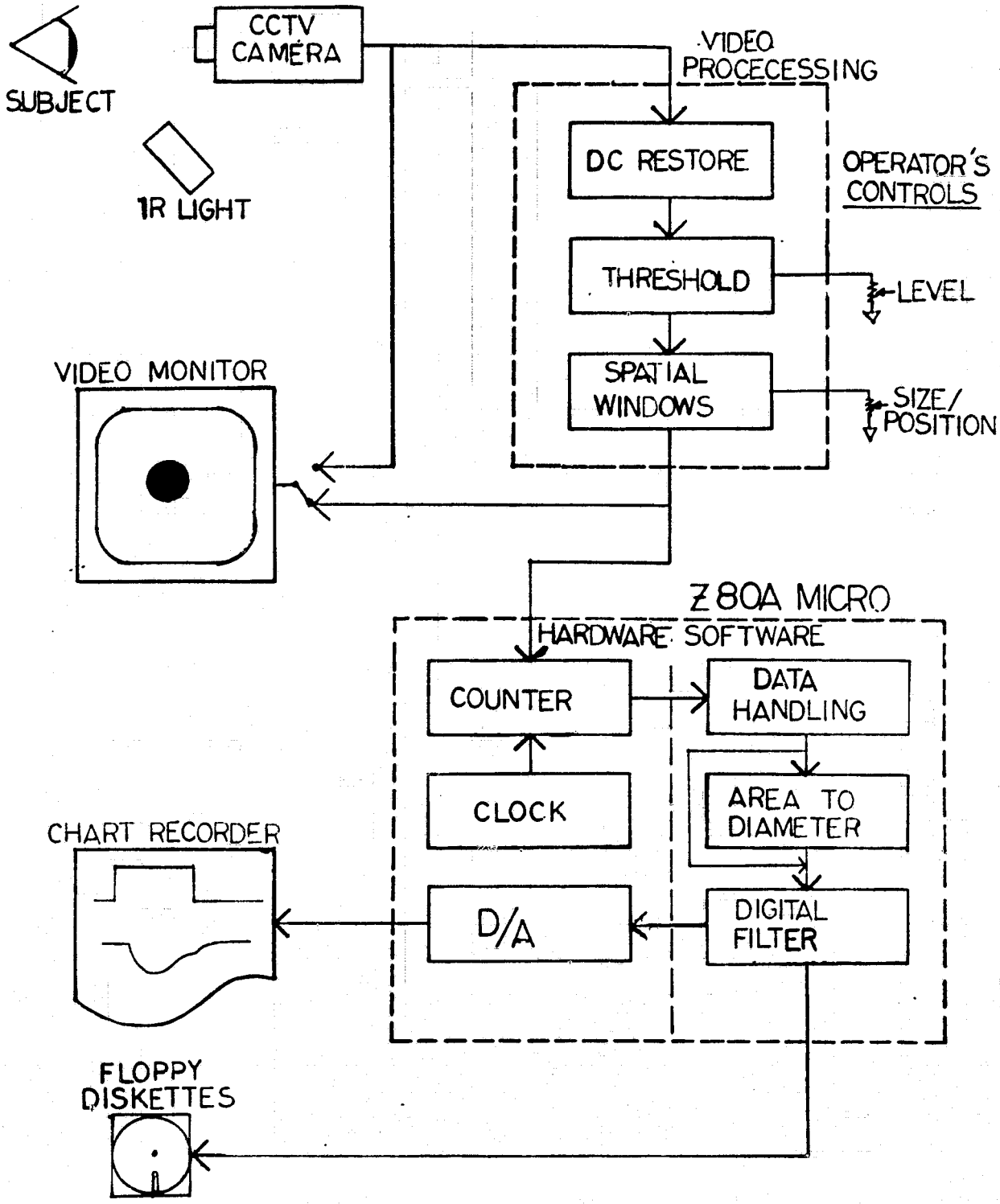


FIGURE 4 MICROPROCESSOR BASED INTEGRATIVE PUPILLOMETER (MIP)

counter to zero for the next frame.

The area may be (optionally) converted to diameter, and output to a digital to analog port. Pupil size (area or diameter) together with other input data (e.g. stimulus) is stored in buffer memory. When the buffer is full, it is automatically dumped to a floppy diskette.

SOFTWARE

The pupillometer was integrated into a microprocessor so as to provide the flexibity of software. Real-time software functions include: Data collection (described above), stimulus generation, area to diameter conversion (planned) . Postprocessing may include correction for eye gaze, correction for optical imperfections, and filtering to remove instrument noise.

PERFORMANCE

The performance of the MIP is summarised in Fig. 5.

SNR (MEASURED)	60 dB
DYNAMIC RANGE	1-9 mm DIAM.
LINEARITY	within .5%
OPTICAL EFFECTS	CORRECTABLE
TEMPORAL BANDWIDTH	30 Hz.
SPACIAL NOISE SENSITIVITY	LOW
INHERENT S/N CHARACTERISTIC (versus diameter methods)	2:1

FIGURE 5 PERFORMANCE OF THE MIP.

EXPERIMENTATION

A. Linear Analysis

It is well known that the pupil control system is nonlinear in several respects. Nevertheless, the techniques of linear systems analysis provide valuable insight into its operation. Fig. 6 shows the system's response to a small-signal sinusoidal light stimulus, provided in Maxwellian view (open loop mode). The response is characterised by the gain and phase of the pupil area signal. Note that open loop gain is defined as:

$$G = \frac{\Delta A_P}{A_P} \Big/ \frac{\Delta L_P}{L_P}$$

where A_P is the area of the pupil

L_P is the light incident on the pupil (and on the retina)

and hence is dimensionless.

If one plots gain and phase for a wide range of frequencies , a Bode plot (Fig. 7) results. Here we see that the pupil system has a high frequency rolloff (third order) at about 2 Hz . Terdiman and Stark (3) showed that this is due to the physiological limitations of the iris, which is a smooth muscle.

The data of Fig. 7 predict that, if the closed loop gain can be raised above 1.0 at about 1.2 Hz, where the phase delay is 180 degrees, oscillations will develop. If the stimulus light is carefully focused on the iris-pupil margin, (see Fig. 8) any change in pupil size results in a large change in the amount of light falling on the retina-i.e. the gain has been greatly increased. When this is done for the subject of Fig. 7, the oscillations shown in Fig. 9 result.

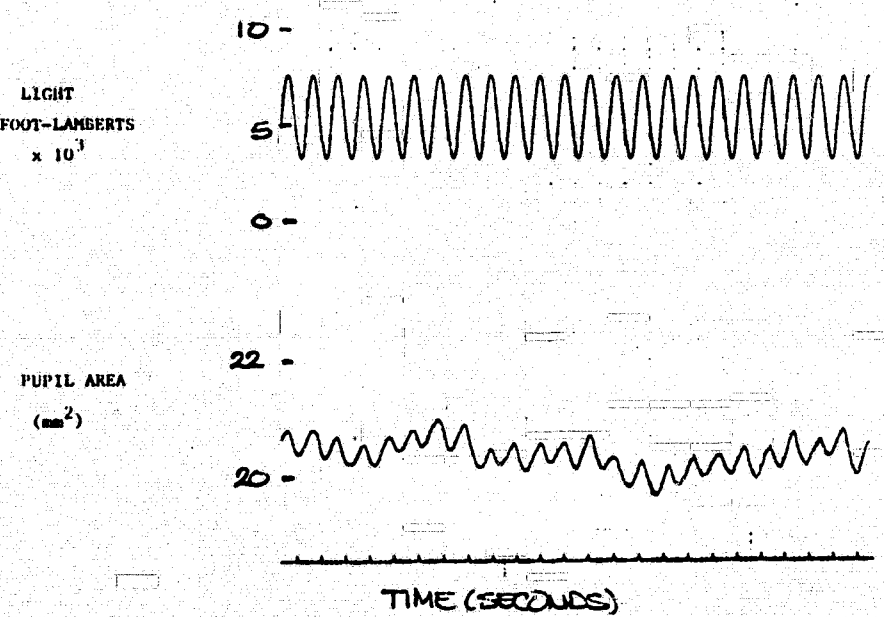
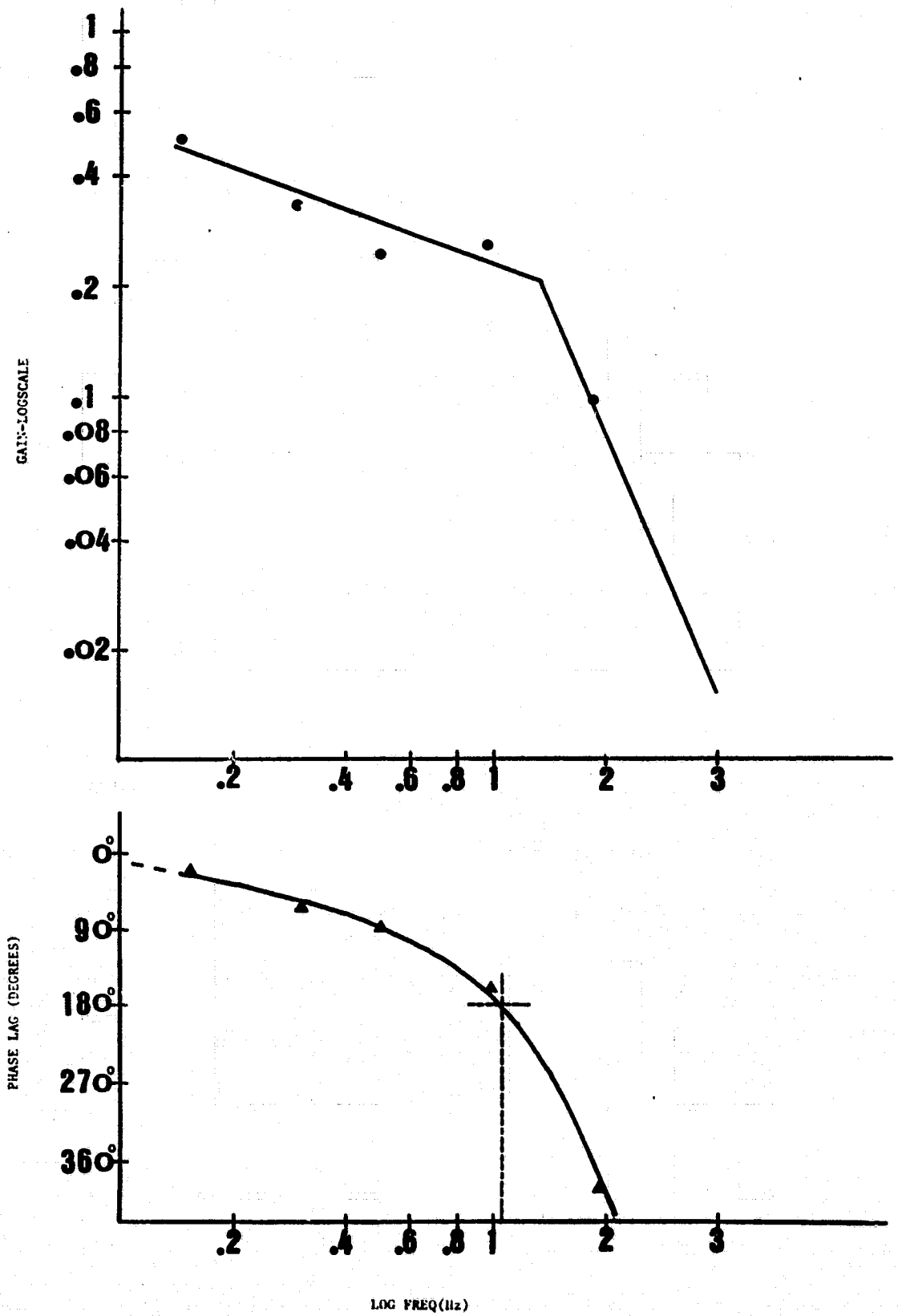


Figure 6 Quasilinear Response to Sinusoidal Light Stimulation.

Frequency= 1.0 Hz. Modualtion Coefficient = 60%

FIGURE 7 BODE PLOT OF PUPIL RESPONSE FOR 1 SUBJECT
VERTICAL LINE REPRESENTS PREDICTED FREQUENCY
OF HIGH-GAIN OSCILLATION(180° PHASE LAG)



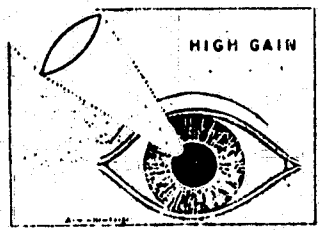


Figure 8 Technique used for stimulation. Light is here focused on border of iris and pupil. Small movements of iris result in large changes in light intensity at retina.

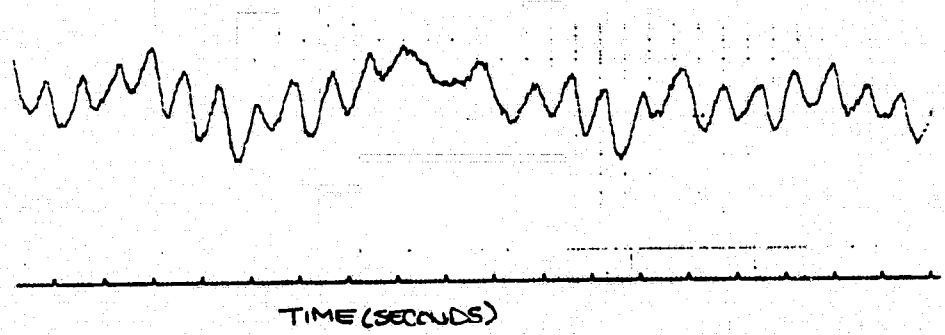


Figure 9 Spontaneous Oscillations Due to High Gain Feedback. Frequency= 1.3 Hz.

B. Nonlinear Phenomena

The pupil exhibits many kinds of nonlinear phenomena. Obviously, it cannot constrict or dialate beyond physiological limits (extremal saturation), more generally, Hansmann, et al have shown that gain is a function of baseline pupil area. It is largest for a medium size pupil , and falls off at the extrema.

Harmonic and sub-harmonic responses may be seen in response to sinsoidal light stimulation. Subharmonics, illustrated in Figure 10, are most evident for large modulation coefficients, and at the system's low frequency limit. E.G. the harmonics in Figure 10 resulted from a 2 Hz sinusoid with 60% modlation. Harmonics are most evident during the low frequency sinusoidal excitation.

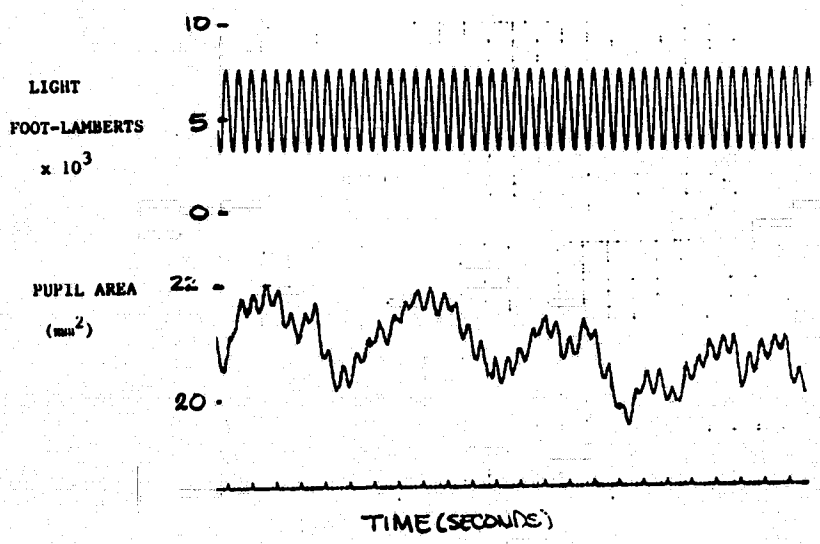


Figure 10 Subharmonic Response to Sinusoidal Light Stimulation

The response to a pulse stimulus (Figure 11) clearly shows a directional assymetery. Regard the pulse as the sum of two step functions, a positive step followed by a negative step. The "on response" is clearly larger than the "off response". In fact, the off response is virtually absent for small pulses. Retinal adaptation allows for pupillary escape, that is for a redilatation of the pupil while the pulse of light is on. This is not a no-memory nonlinearity, and has to be modelled as a frequency dependent gain change.

SIMULATION

Shimizu developed a computer model of the pupil to simualate the effects illustrated in Figures 6 thru 11. A block diagram of the model is shown in Figure 13, and the corresponding results are shown in Figure 12

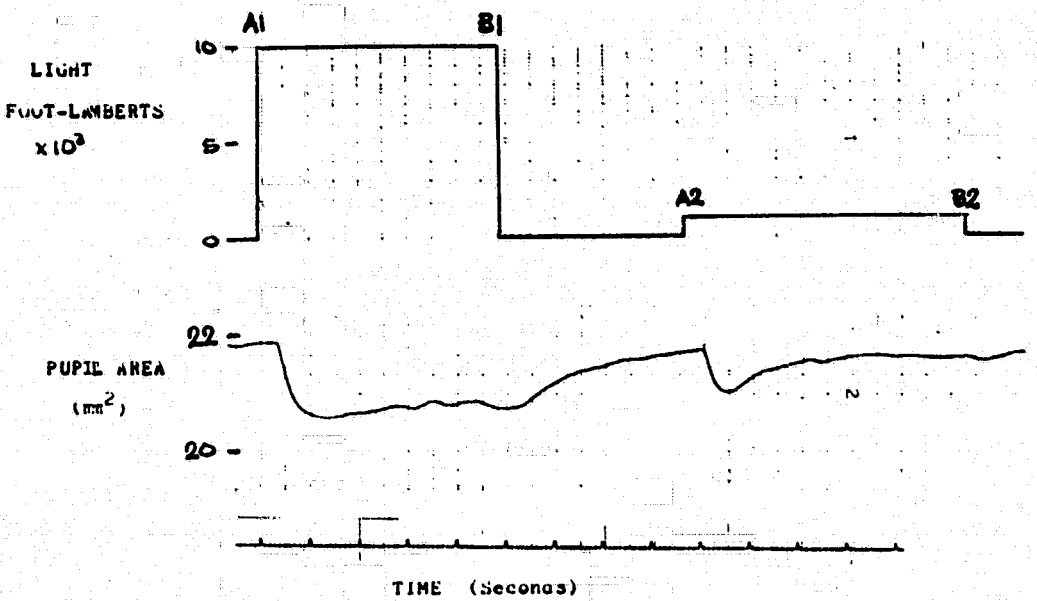


FIGURE 11. RESPONSE TO PULSES OF LIGHT SHOWING DIRECTIONAL ASYMMETRY, DIFFERENT TIME CONSTANTS OF RESPONSE, AND PUPILLARY ESCAPE

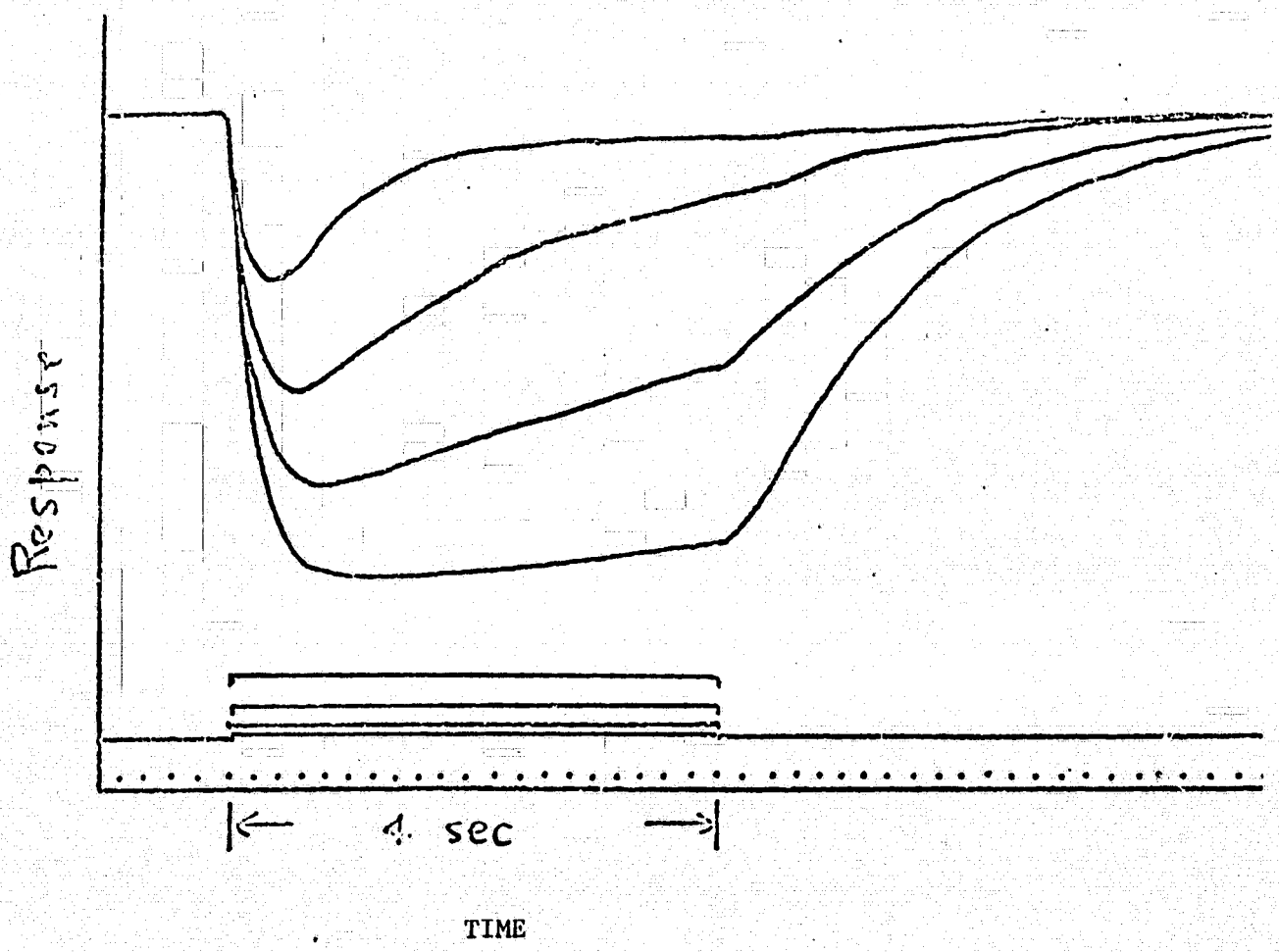


FIGURE 12. PUPILLARY ESCAPE SIMULATION RESULTS

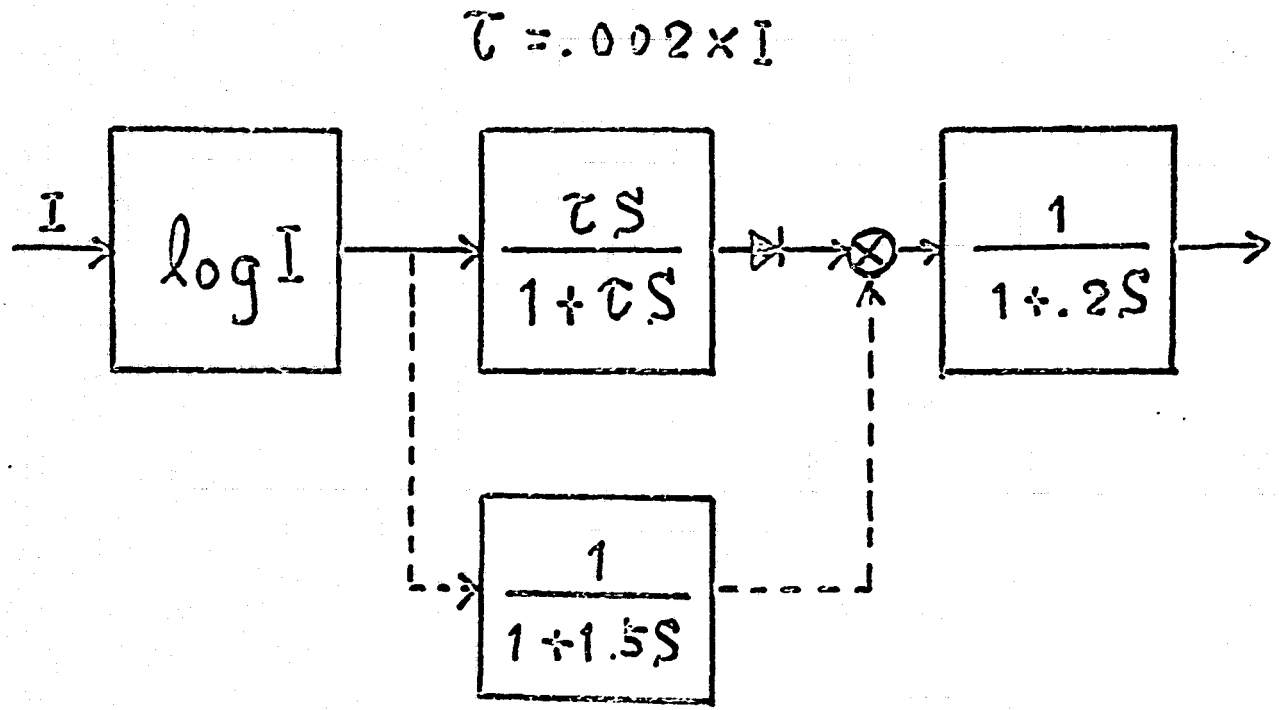


FIGURE 13. PUPILLARY ESCAPE MODEL

Acknowledgement

We are pleased to acknowledge the partial support from the NCC 2-86 Cooperative Agreement, NASA Ames Research Center, and from the NIH Systems and Integrative Biology Training Grant.

REFERENCES

1. Lowenstien, O. and Lowenfeld, I.E., A.M.A. Archiv. Opthal. vol 59, no. 3, 352-363 (1958)

2. Stark, L.,P.I.R.E. vol 47, (1959)

3. Terdiman J., Smith J.D., and Stark L. Brian Research vol 16 (1969) pp288-292.

4. Hansmann D., Semmelow J., and Stark L. (1974) A Physiological Basis of Pupil Dynamics. In: Pupillary Dynamics and Behavior. (Ed. M.P. Janisse) Plenum, New York.

D47

MODEL SIMULATION STUDIES TO CLARIFY THE EFFECT ON SACCADIC EYE MOVEMENTS OF INITIAL CONDITION VELOCITIES SET BY THE VESTIBULAR OCULAR REFLEX (VOR)

By Moon-hyon Nam, Jack M. Winters and Lawrence Stark

University of California, Berkeley

Acknowledgement: We are pleased to acknowledge partial support from the NCC 2-86 Cooperative Agreement, NASA-Ames Research Center and the Ministry of Education, Korea.

SUMMARY

Voluntary active head rotations produced vestibulo-ocular reflex eye movements (VOR) with the subject viewing a fixation target. When this target jumped, the size of the refixation saccades were a function of the ongoing initial velocity of the eye. Saccades made against the VOR were strikingly attenuated in magnitude while saccades going with the VOR were larger in magnitude. Simulation of a reciprocally innervated model eye movement provided results comparable to the experimental data.

Most of the experimental effect appeared to be due to linear summation for saccades of 5 and 10 degree magnitude. For small saccades of 2.5 degrees, peripheral nonlinear interaction of state variables in the neuromuscular plant also played a role as proven by comparable behavior in the simulated model with known controller signals.

INTRODUCTION

Under natural conditions, it is well known that different types of eye movements, generated by specialized subsystems of the CNS, may interact temporally in a complex manner. In recent years, the interactions between different types of eye movements has been studied in humans and primates in a number of different ways and the linear summation theory for such movements has been proposed. In particular, data on voluntary saccades with compensatory movements (VOR) in monkeys (Morasso et al (1973)), saccade/vergence interactions in man (Kenyon et al (1980)); Ono et al(1978)), goal-directed saccades and VOR in man (Jurgens et al (1981)) all support the theory of additivity when two types of eye movements combine. Chun and Robinson (1978) postulated switching off of slow command during the execution of quick phase VOR in monkey. Also, earlier work by Nam et al (1981) presented some evidence for the 'Kenyon effect' in saccade and VOR interaction.

In this study, we examined our former hypothesis on linear summation through the interaction between saccade and VOR eye movements by experiments in humans and by computer simulation using a modified version of the eye model described previously by Lehman and Stark (1979). Were the observed phenomenon due to 1) linear summation of VOR with the saccades, 2) peripheral nonlinear interaction of state variables in the neuromuscular plant, i.e., the 'Kenyon effect', 3) nonlinear interaction at ocular motor neurons, the final common path that might have shown nonlinear squelching phenomena, or 4) higher level preprogrammed changes in saccadic magnitude? By combining experimental results with simulation findings, in which the latter has a known controller signal, cases 1) and

2) can be separated from cases 3) and 4), and significant insights can be obtained.

METHOD

A. Experimental

Horizontal eye movements were measured using the infrared photoelectric method which has a bandwidth of 1 kHz and a sensitivity of 10 min of arc. Measurements were linear for a range of $\pm$ 25 degrees and frequent calibrations guaranteed that recorded eye movements faithfully reflected retinal orientation.

Horizontal head movements were measured with an electro-mechanical transducer system consisting of mechanical linkages using light-weight universal joints and a sliding mechanism coupled to a low torque special film potentiometer. The linkage was self-aligning with the vertical axis of the head and thus only rotational motion of the subject's head in the horizontal plane was measured by the potentiometer. This system allowed for flexible and natural head movements for the subject.

Data was obtained for five adult subjects. To generate the appropriate VOR, the subject would rotate his head voluntarily at a frequency between 0.8 and 1.5 Hz in a sinusoidal-like motion with a peak amplitude between $\pm$ 10 and $\pm$ 15 degrees relative to the center 'straight-ahead' position. Calibrations were performed on eye movements within $\pm$ 25 degrees of the visual target using 5 equidistant points at a distance of 19 cm from the eye axis of rotation. After separate calibrations for the eye and head movements to the same calibration target, the task of the subject was to follow the jumping targets as quickly as possible while making the sinusoid-like head rotations.

The experimenter instructed the subject to carefully follow the visual target and also gave the subject's various head rotation schemes (eg. faster, slower, larger, smaller so as to obtain data with a variety of head rotations), all the time viewing the eye and head movement recordings. Vision was monocular to avoid vergence effects. The data was classified in terms of the initial velocity conditions as being 'with' (W) or 'against' (A) the VOR.

B. Simulation

The 'core' eye model algorithm has been described in detail previously (Lehman and Stark, 1979). The corresponding program represented

a simulation of the eye mechanics for horizontal eye movements, i.e. of the lateral and medial rectus muscle, involves the integration of a set of six nonlinear differential equations.

The present program represents a significant extension of the model so as to allow a sinusoid 'VOR' neurological signal---of specified magnitude and frequency---to be the 'base' control signal. A saccade (second order, of 2.5, 5, 10 or 20 degrees magnitude) can then be produced at any time during a run, and consequently for different initial conditions of VOR position, velocity, and acceleration. As presently set up, there is both a left and a right saccade, with the left occurring first.

The main outputs of the program are the eye position and velocity as a function of time. In addition, a number of relevant data parameters are calculated directly by the program, including the initial position and velocity of the system at the start of the saccade, the apparent and relative peak saccadic velocities, and the duration and magnitude of the resulting saccade (based on 'returning velocity' considerations).

RESULTS

Relative velocity distinguished from apparent velocity. Before presenting the primary results, it is important to distinguish between the apparent peak velocity, which is defined as the velocity with respect to the absolute zero velocity, and the relative peak velocity, i.e., the peak velocity relative to the initial velocity at the start of the saccade. The difference between the two, which is due to the velocity of the compensatory eye movement when the saccade begins, can be up to approximately 100 deg/sec, which is obviously significant relative to typical peak saccadic velocities. As a limiting case, consider the difference between the absolute and relative peak velocities for the small 'correction' saccades (Figures 1 and 2). In particular, notice that for the right-most saccade in Figure 2, which has an amplitude of about 1 degree, the 'apparent' peak velocity is -30 deg/sec, while the 'relative' peak velocity is about 90 deg/sec---quite a difference. Clearly, 90 deg/sec is the appropriate velocity for this positive going, leftward saccade. This difference, of course, is more noticeable for small saccades. Still, it is obvious that there will be significant difference between figures using the two different peak velocity definitions. To see this more explicitly, consider a typical simulation run (Figure 3). Notice that the 'against' saccade, which by definition has a negative initial velocity, has a proportionally smaller 'absolute' versus 'relative' peak velocity. Conversely, for the 'with' saccade, the 'absolute' peak velocity is proportionally larger. Notice also that there is a linear relationship between the initial velocity and the peak velocity measurements. Thus, an 'absolute' peak velocity versus

initial velocity curve can be turned into a similar 'relative' graph just by adjusting the slope in an appropriate manner.

Main effect of VOR velocity as an initial condition. The main effects of the VOR initial conditions can be seen in a qualitative manner (Figures 1, 2 and 3). The magnitudes of the saccades are strikingly different depending upon whether they are on or against the initial condition velocity set by the VOR (Figure 1). The saccades against are attenuated and the saccades with are increased. It is interesting to note that saccades against are attenuated with respect to saccades where the initial condition is at 0 degree velocity (Figure 2). Similar results are found in the simulation traces (Figure 3). These similarities between experimental and simulation traces will be shown in a more qualitative fashion in later figures. Since the simulation results were obtained with a known controller signal (CS), by comparing simulation versus experimental results insights can be gained into the mechanism(s) responsible for the observed eye interaction phenomena. Magnitude of a large number of experimental and simulated saccades following 10 degree target jump can be plotted as a function of the initial eye velocity of the VOR (Figure 4). In addition to the experimental and simulation curves, a line representing the projected system behavior if there was ideal linear summation of the two controller signals is also shown. Notice the strong correlation between the experimental and simulation data and the linear summation line. This provides very strong evidence for the linear summation hypotheses. A similarly strong correlation also existed for 5 degree target saccades (not illustrated).

Hysteresis. A secondary observation was that, in the simulation results, a 'hysteresis' loop is apparent. This 'hysteresis' behavior was due to the initial position and acceleration conditions at the start of the saccade, which should be taken as implicit parameters in the displayed curve. When a 'with' saccade is past the center 'zero' position of the eye, the initial position and acceleration conditions will both tend to oppose eye motion, the former by virtue of the system elasticity and the latter by virtue of system inertia and its resulting deceleration. Thus, this could be called a 'with-against' case. By extension there are also 'with-with', 'against-with', and 'against-against' cases. These simulation results indicate that some of the experimental scatter, which is already not very large, may be due to the initial position and acceleration conditions which are implicit parameters in the graph.

Saccadic magnitude as a function of initial condition velocity and of intended saccadic size. Averaged experimental data for the four different nominal saccade sizes is compared with the projected linear summation case (as solid line) (Figure 5). Notice that, in general, the slope of the experimental data is slightly lower. Part of this difference could be

directly due to the fact that average VOR velocity would be a better representation than the initial velocity used. Normalized magnitude is plotted versus the target magnitude so as to compare the experimental magnitude data with the expected data based on linear summation for initial conditions of 40 and 80 degrees per second (Figure 6). Again we see that linear summation accounts for a bulk of the effects. However, experimental results show less effect than that expected from linear summation. This could be due to the assumption that initial condition velocity lasts throughout the saccade; there might be changes during the saccade that would alter the summation calculations.

Saccadic velocity as a factor of initial condition velocity. Additional evidence for linear summation can be found through the analysis of initial condition velocity and apparent peak velocity relationships. Experimental data were compared with data from simulation using the hypothetical linear summation model (Figure 7). The simulation results also show the 'hystereses' effects of initial eye position and acceleration (also seen in Figure 4). Notice that the apparent peak velocity is graphed versus the initial VOR velocity. For reasons explained above, the relative peak velocity could also have been plotted. In the latter case the linear summation hypothesis would give a horizontal line. Naturally, the experimental and simulation data would change in a corresponding manner. Differences between the experimental and simulation data were partly due to this difference of reference.

Main sequence relationships. The relationship between the saccadic magnitude, duration and apparent peak velocity (Figure 8) shows further evidence of interaction. Experimental and simulation data on apparent velocity versus magnitude show different relative slopes for each saccade size, crossing the normal main sequence (solid line) with a significant shift in some cases. The amount of shift between the apparent and relative peak velocity, which corresponds to the initial condition velocity, could explain some of this shift. Mean values and 'with' and 'against' for 5 and 10 degree saccades fall on the main sequence. For 2.5 and 20 degree saccades, while the mean values fall on the main sequence, the with and against initial condition velocity set by VOR lie off the main sequence. The reason for the shift away from the main sequence might be non-main sequence effects of VOR velocity and indicates there are non-linear phenomena with these very small and very large saccade sizes.

REFERENCES

Morasso, P., Bizzi, E., and Dichgans, J. Adjustment of Saccade Characteristics during Head Movement. Exp. Brain Res., 16: 492-500, 1973.

Kenyon, R.V., Ciuffreda, K.J. and Stark, L. Unequal Saccades during Vergence. Am. J. Optom. & Physio. Opt. 57(9): 586-594, 1980.

Ono, H., Nakamizo, S., and Steinbach, M.J. Nonadditivity of Vergence and Saccadic Eye Movement. Vision Res., 18: 735-739, 1978.

Jurgens, R., Becker, W., Rieger, P., and Widderich, A. Interaction between Goal-directed Saccades and the Vestibulo-ocular Reflex (VOR) is Different from Interaction between Quick Phases and the VOR. In Progress in Oculomotor Research, ed. by Fuchs, A. and Becker, W., Elsevier, North Holland, #33, 1981.

Chun, K.S. and Robinson, D.A. A Model of Quick Phase Generation in the Vestibulo-ocular Reflex. Biol. Cyber. 28: 209-221, 1978.

Nam, M., Winters, J. and Stark, L. Simulation and Nonlinear Interaction of VOR with Saccadic Eye Movements. OMS 81, Jan. 18-21, 1981.

Lehman, S. and Stark, L. Simulation Analyses and Enumeration Studies of Time Optimal Control. J. Cyber. Inf. Sci. 2(2/4) 21-43, 1979.

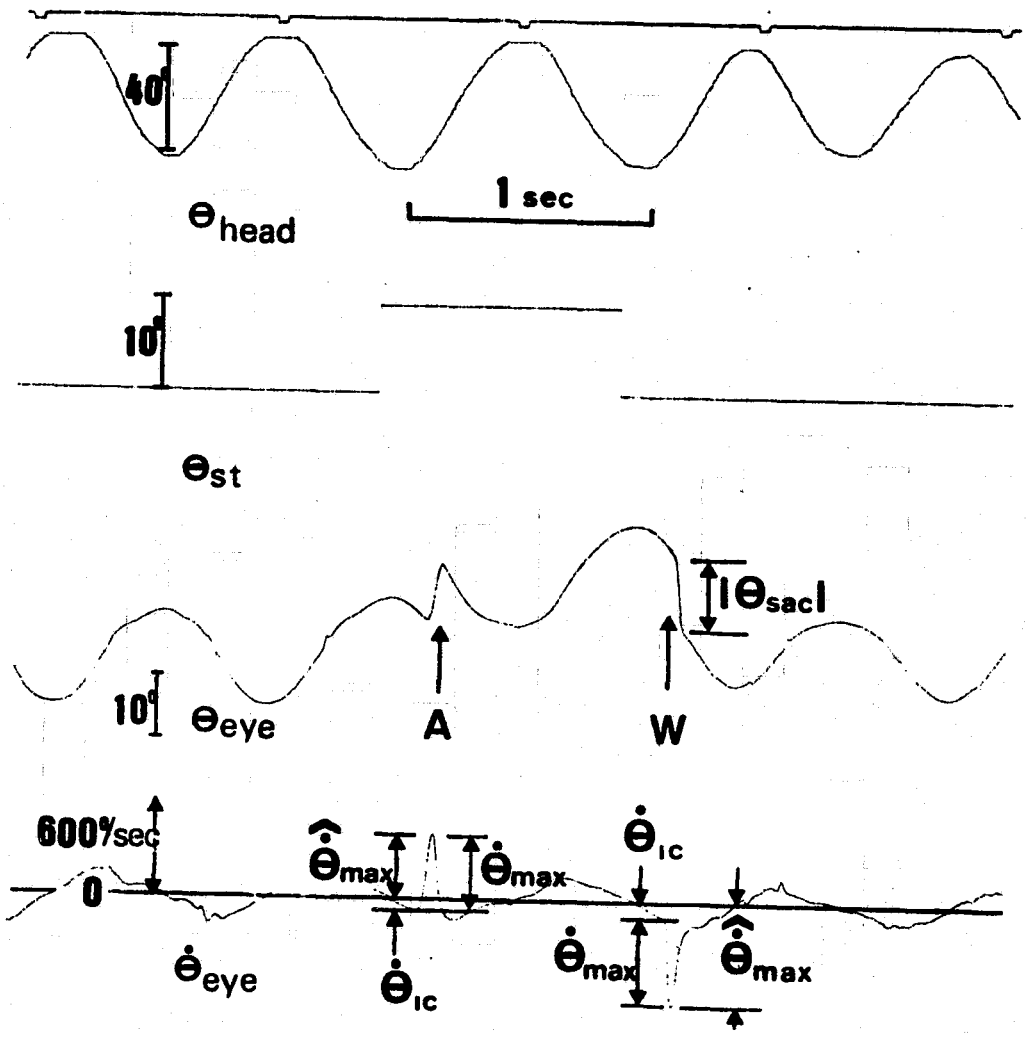


Figure 1. Recordings of head position (θ_{head}), visual target (θ_{st}), eye position (θ_{eye}) and eye velocity ($\dot{\theta}_{eye}$). In eye position trace, A and W stand for against and with VOR, and saccadic magnitude ($|\theta_{sac}|$) and duration are measured from saccadic portion. In eye velocity record, $\dot{\theta}_{ic}$ = initial velocity of VOR eye movements at the instant of saccadic occurrence, $\hat{\dot{\theta}}_{max}$ = apparent peak velocity.

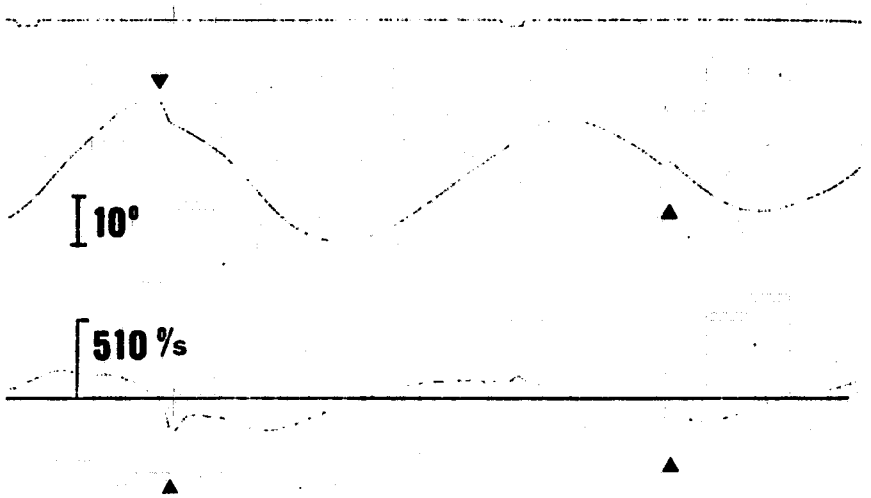


Figure 2. Eye position and velocity record illustrating apparent and relative peak velocity.

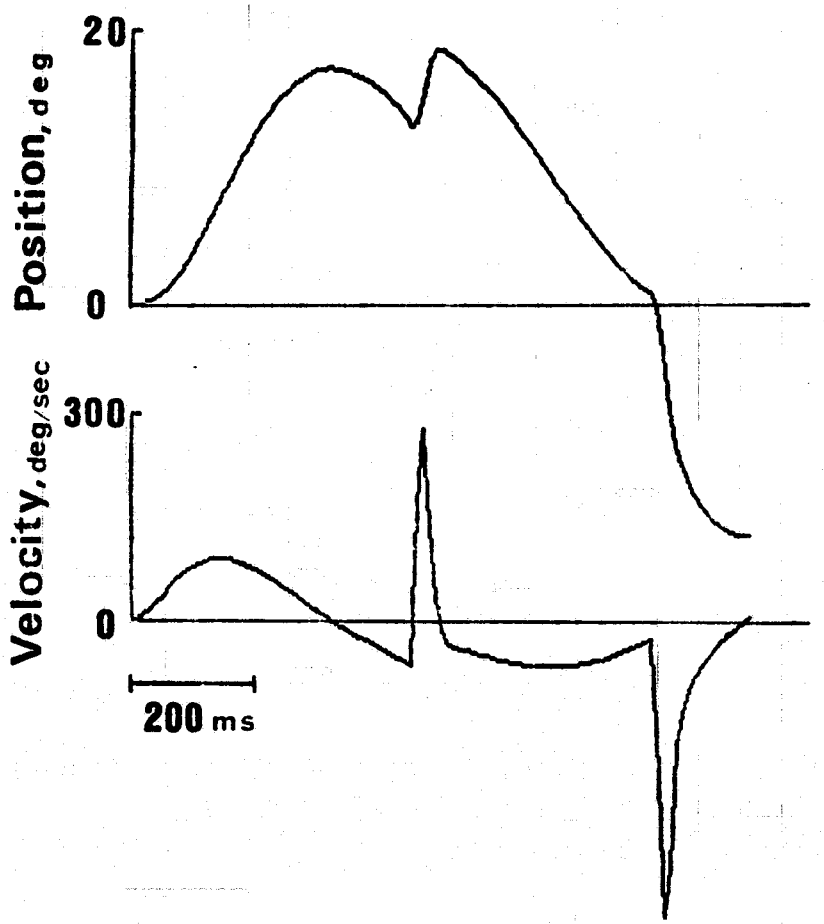


Figure 3. Computer simulation trace for 5 degree saccade and 1 Hz, 35 degree sinusoid VOR controller signal.

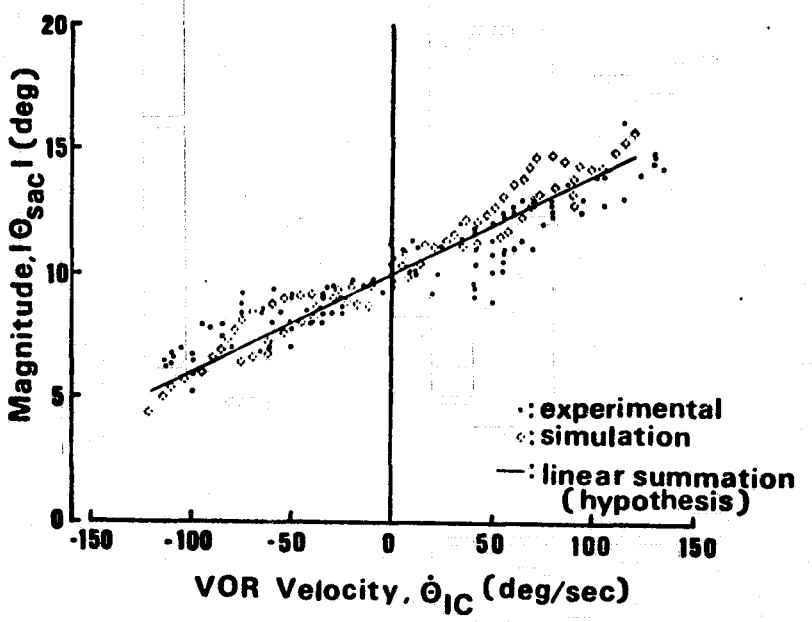


Figure 4. Magnitude vs initial condition velocity relationship for 10 degree target.

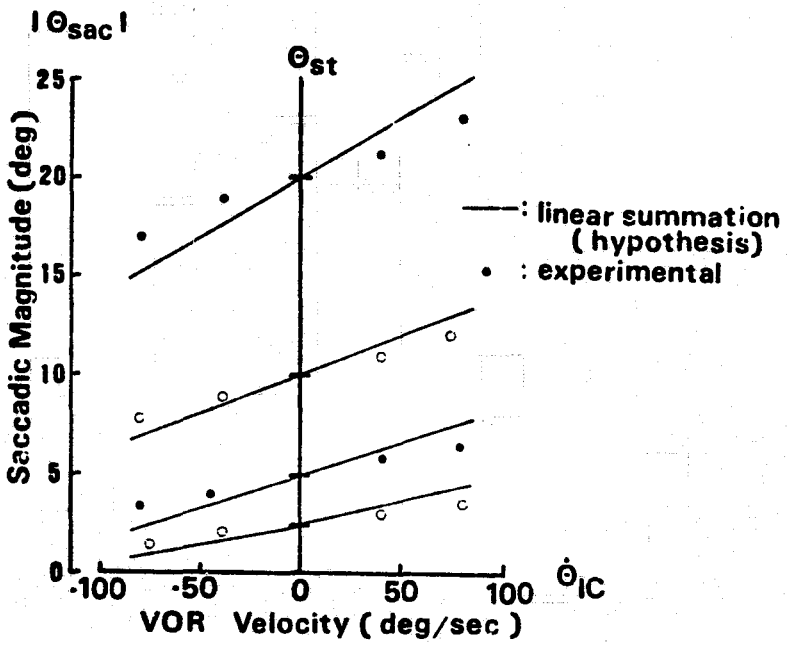


Figure 5. Averaged magnitude vs initial condition velocity relationship for four nominal target jump.

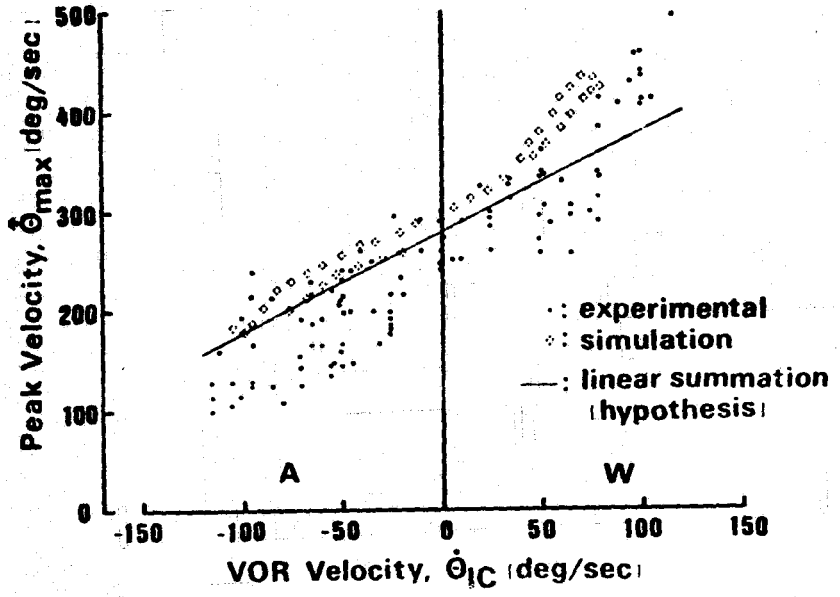


Figure 6. Normalized magnitude vs four nominal target.

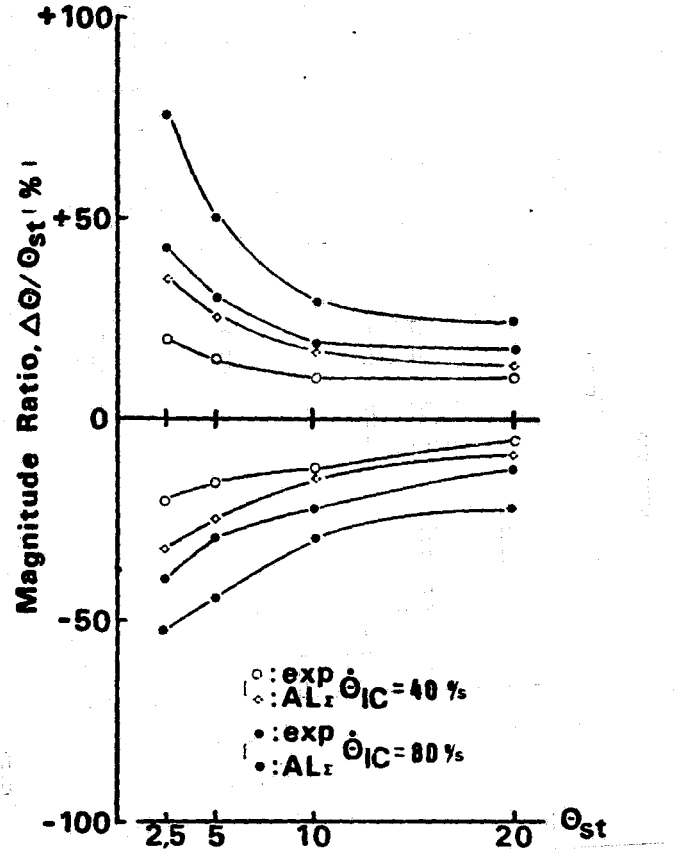


Figure 7. Apparent peak velocity vs initial condition velocity relationship for 5 degree target.

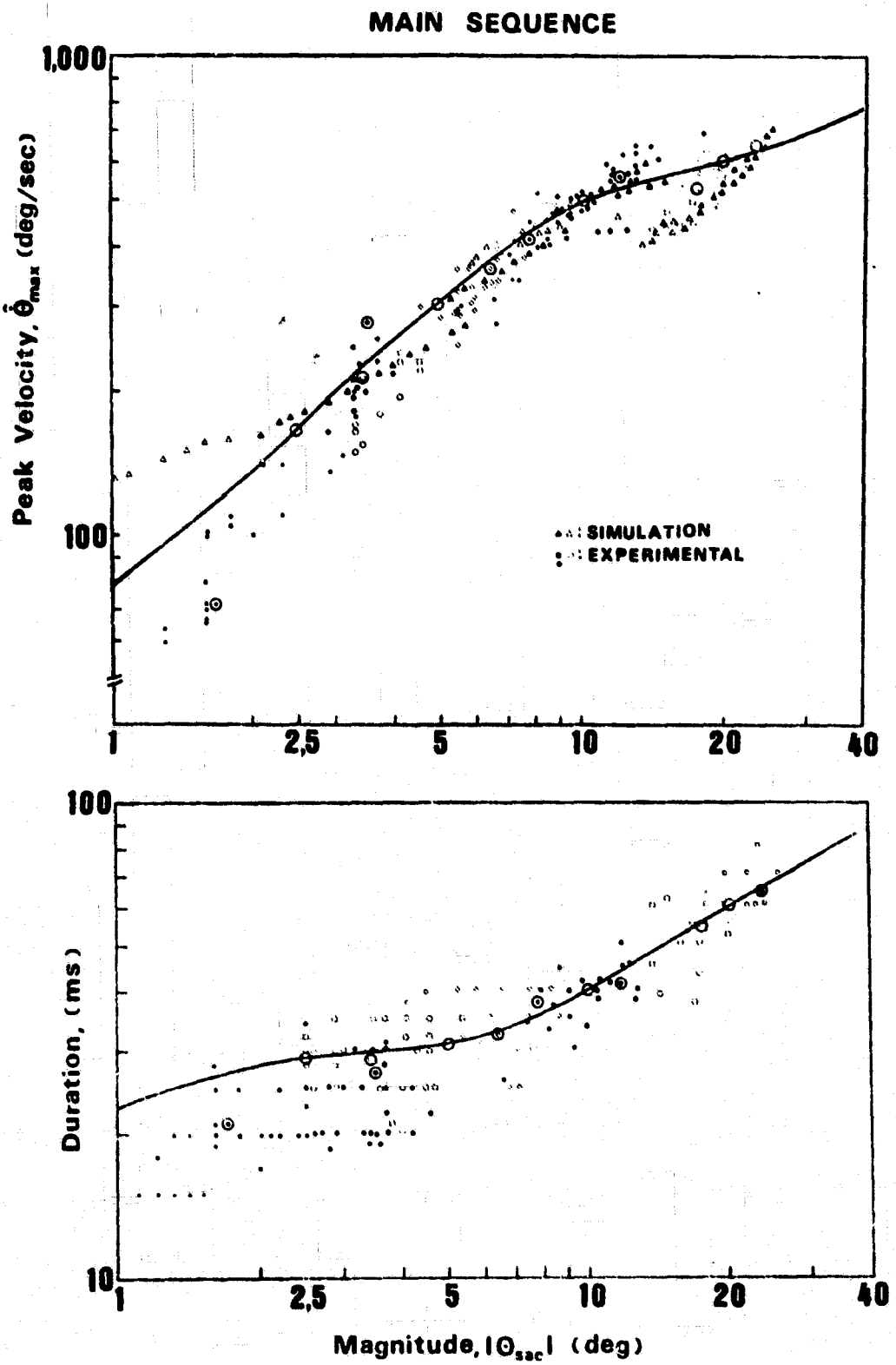


Figure 8. Main sequence relationship for apparent peak velocity, duration vs magnitude.

D48

BENEFITS OF DETAILED MODELS OF MUSCLE ACTIVATION AND MECHANICS

By Steven L. Lehman and Lawrence Stark

University of California, Berkeley

SUMMARY

Recent biophysical and physiological studies have identified some of the detailed mechanisms involved in excitation-contraction coupling, muscle contraction, and deactivation. Mathematical models incorporating these mechanisms allow independent estimates of key parameters, direct interplay between basic muscle research and the study of motor control, and realistic model behaviors, some of which are not accessible to previous, simpler, models. The existence of previously unmodeled behaviors has important implications for strategies of motor control and identification of neural signals. New developments in the analysis of differential equations make the more detailed models feasible for simulation in realistic experimental situations.

INTRODUCTION

Mathematical models and computer simulations are often used in manual control studies in an attempt to deduce the properties and strategies of the controller from the dynamical behavior of the whole system. In such inverse problems, the properties of the plant (muscles and load) must be carefully identified in order for the deduced model input to reflect the actual system control signal.

This identification problem for neuromuscular systems has attracted the attention of two groups of investigators, with two divergent points of view (reference 1). Biomedical engineers have tended to construct and identify models on the basis of macroscopic mechanical behavior, while muscle physiologists and biophysicists have concentrated on detailed microscopic mechanisms.

While not all microscopic mechanisms have macroscopically significant influences, i.e., unobservable states, some profoundly affect observable behaviors. We present two examples, one in muscle mechanics and one in activation/deactivation, for which detailed biophysical models have distinct advantages over the more widely used phenomenological ones.

MUSCLE MECHANICS

The dichotomy between macroscopic phenomenological description and microscopic mechanism is clear in the two prevalent classes of models of muscle mechanics. Engineers tend to use the classic three-element model (figure 1), while muscle physiologists consider ever more detailed cross-bridge models (figure 2). It is instructive to compare the two types with respect to the three main mechanical characteristics of muscle: the static length-tension relationship, the force-velocity curve, and the transient behavior evident in quick length change experiments.

Length-Tension Curve

The length-tension characteristic for passive muscle is of course independent of the contractile mechanism per se, so is modeled the same way for both types (element labelled PE in figure 1). The characteristic added for active muscle, on the other hand, was not explained until Gordon et al (reference 2) invoked a cross-bridge model, and showed that the active characteristic was simply the result of varying cross-bridge overlap. Cross-bridge models thus have the advantage of a natural implementation of the length-tension curve, and comparability with an actual measurement (the filament lengths as measured from electron micrographs).

Although the classical phenomenological model neither explains the full length-tension curve nor implements it elegantly, it may be made to exhibit the known characteristic. In fact, the length-tension curve is generally included in this model ad hoc as an additional, length-dependent element.

Force-Velocity Relationship

By the force-velocity relationship we mean both the relationship between force and (constant velocity) shortening velocity first characterized by Fenn and by A.V. Hill (reference 3) and its extension to steady-state force exerted by a muscle lengthening at constant velocity (figure 3). Hill fit the shortening curve with his well-known hyperbola, the two parameters of which he related to the maximum shortening velocity and the shortening heat.

The force-velocity relationship has been included in the phenomenological models in various ways, both directly (as part of the box labelled CE in figure 1) and as a velocity-dependent viscosity. The Hill formalism makes it possible to construct the entire shortening characteristic from two constants---a compression of experimental data valuable in computation.

The force-velocity relationship (both shortening and lengthening) is produced by the cross-bridge models, as a consequence of the choice of kinetic rate constants between cross-bridge states. Here again the cross-bridge models are more elegant than the phenomenological type, because they explain the observed macroscopic effect from a lower level. The production of the force-velocity curve is not, however, surprising. Indeed, the rate constants are chosen to fit the curve. A.F. Huxley explained the relationship between his rate constants and the Hill constants in his report of the first cross-bridge model. (reference 4)

The most significant difference between the two types of model is not shown in Figure 3. It is now well-known that the curve for lengthening muscles (velocity less than zero in the figure) is only valid in the steady state. Actual muscle, when lengthened at constant velocity, produces first more force than that indicated by the curve, then yields to tensions lower than those indicated (references 5,6). This transient behavior of lengthening muscle is not produced by the phenomenological models, because their imposed force-velocity relationships are single-valued. This behavior is produced by almost any cross-bridge model, including the simplest (two-state) models.

Transients

The series elastic element (SE in figure 1) was introduced into the phenomenological models to account for the changes in muscle tension observed during quick stretches and releases. It was observed that the response consisted of at least two phases, the first of which implied the existence of an elasticity in series with the active contractile machinery.

Clearly, the addition of a series elasticity makes an allowance for this important compliance, but does not fully solve the problem of the transient. The simulation of the quick-stretch and quick-release data in detail is possible using cross-bridge models.

ACTIVATION AND DEACTIVATION

The complex of processes comprising muscle activation, from the arrival of a nerve action potential to the binding of myosin heads to actin, and the process of deactivation by active pumping of calcium into the sarcoplasmic reticulum have been intensively studied in recent years. Corresponding to the increase in understanding of these fundamental mechanisms, an immensely rich, complex, and frustratingly fragmented literature has grown up. The synthesis of this literature and building of models will certainly provide better estimates of the time scales and relative influences of the many processes involved, and may also reveal new dynamical possibilities.

Already, there are specific, well-identified models for many of the individual processes. For example;

1. Active invasion of the T-system by action potentials has been modeled. (Adrian and Peachey, reference 7)
2. A gating mechanism for calcium release from the sarcoplasmic reticulum has been shown, and its voltage dependence found. (Schneider et al, reference 8)
3. The detailed biochemical kinetics of the protein that pumps calcium into the sarcoplasmic reticulum, thus relaxing muscle, have been investigated, to the level of finding twelve distinct biochemical states of the enzyme and rate constants between those states. (Inesi, reference 9)

The extreme reductionism of the muscle activation and deactivation studies has both good and bad effects. The unfortunate fragmentation and specificity of the large literature inhibits synthesis of results and evaluation of the relative importance of different effects. On the other hand, the reductionistic trend means that the mechanisms found are characteristic of specific proteins, for example, and not of specific muscles or organisms. Because these proteins are likely to be used in all sorts of muscles, the models may be more generally useful. For example, the calcium-pumping protein mentioned above seems to have the same kinetic properties in many types of vertebrate striated muscle. Furthermore, its concentration in sarcoplasmic reticulum membrane is very nearly constant. Therefore, from estimates of the surface area of the sarcoplasmic reticulum easily obtained from electron micrographs, one can reduce the general time course of deactivation of the muscle. Such a conclusion is exceedingly difficult to draw from other (e.g. dynamical) data. (reference 10)

CONCLUSIONS

There are several clear advantages to using detailed biophysical models for muscle activation, deactivation and mechanics. Among them:

1. Such models allow direct comparison with basic muscle research.
2. Some of the detailed models have <u>behaviors</u> that are not in the repertoire of simpler, phenomenological models:

2a. Yielding in strongly stretched lengthening muscle.

2b. Dependence of time constants of deactivation on history that allows for a fused tetanus at lower tonic firing rates.

3. Biophysical models allow independent estimation of mechanically influential parameters from simple measurements (e.g., time constant of deactivation from measurements of electron micrographs.)
4. Detailed mechanistic models permit the natural inclusion of known characteristics (e.g., the length-tension curve for active muscle.)

The <u>disadvantages</u> of such detailed models are, of course, clear to bio-engineers. Conceptual and computational difficulty are the main ones;

these are, of course, diminishing with increasing specialization of humans and power of computers. While some detailed mechanisms have important influences on macroscopic observables, others do not justify their computational cost for manual control studies. The advantages listed above are, however, compelling reasons for the consideration of more detailed and biophysical models.

REFERENCES

1. Hatze, H. , Neuromusculoskeletal Control Systems Modeling -- A Critical Survey of Recent Developments, IEEE Trans. Auto. Control, AC25, 3, 1980.

2. Gordon,A.M., Huxley,A.F., Julian,F.J., The variation in isometric tension with sarcomere length in vertebrate muscle fibres, J. Physiol., 184, 170-192, 1966.

3. Hill, A.V., The heat of shortening and the dynamic constants of muscle, Proc. Roy. Soc.B, 126, 136-195, 1938.

4. Huxley,A.F., Muscle structure and theories of contraction, Progr. Biophys., 7, 255-318, 1959.

5. Rack,P.M.H., The behavior of a mammalian muscle during sinusoidal stretching, J.Physiol., 183, 1-14, 1966.

6. Rack, P.M.H., Westbury,D.R., The short range stiffness of active mammalian muscle and its effect on mechanical properties, J.Physiol, 240, 331-350, 1974.

7. Adrian,R.H., Peachey,L..D., Reconstruction of the action potential of frog sartorius muscle, J.Physiol., 235, 103-131, 1973.

8. Schneider,M.F.,Chandler,W.K., Vol age-dependent charge movement in skeletal muscle: a possible step in excitation-contraction coupling, Nature, 242, 244-246, 1973.

9. Inesi,G., Kurzmack,M., Coan,C., Lewis,D.E., Cooperative calcium binding and ATPase activation in sa coplasmic reticulum vesicles, J.Biol.Chem, 1980.

10. Lehman,S.L., Ph.D. Thesis, University of California, Berkeley, 1981.

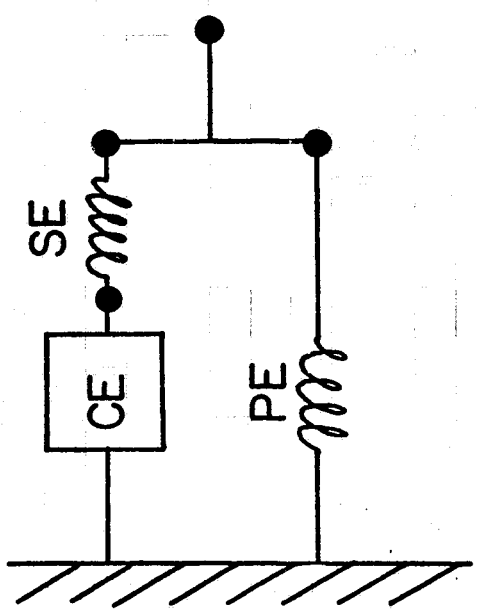
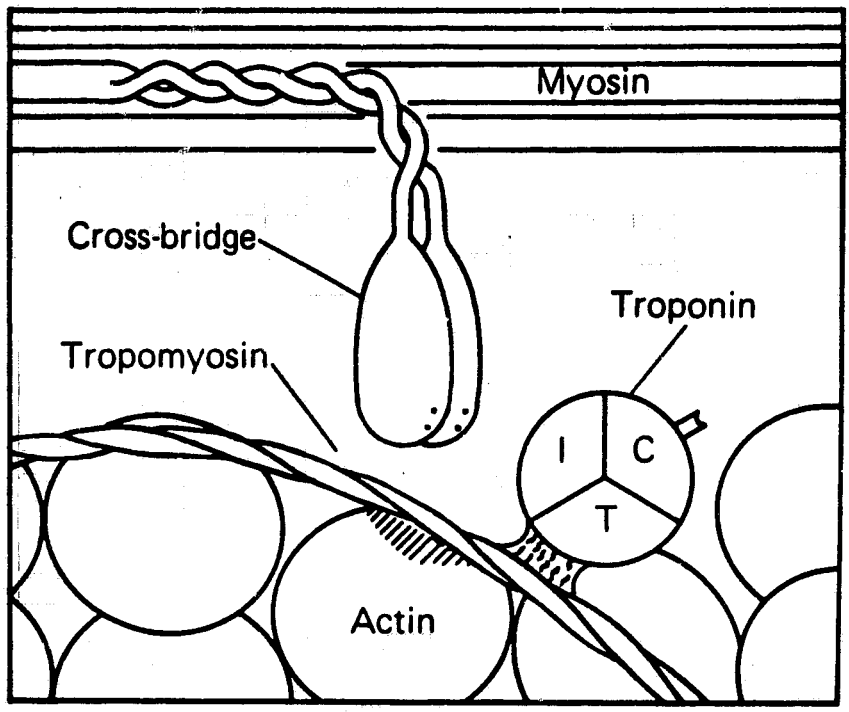


Figure 1: Classical 3-element muscle model
CE : contractile element
PE : parallel elasticity
SE : series elasticity



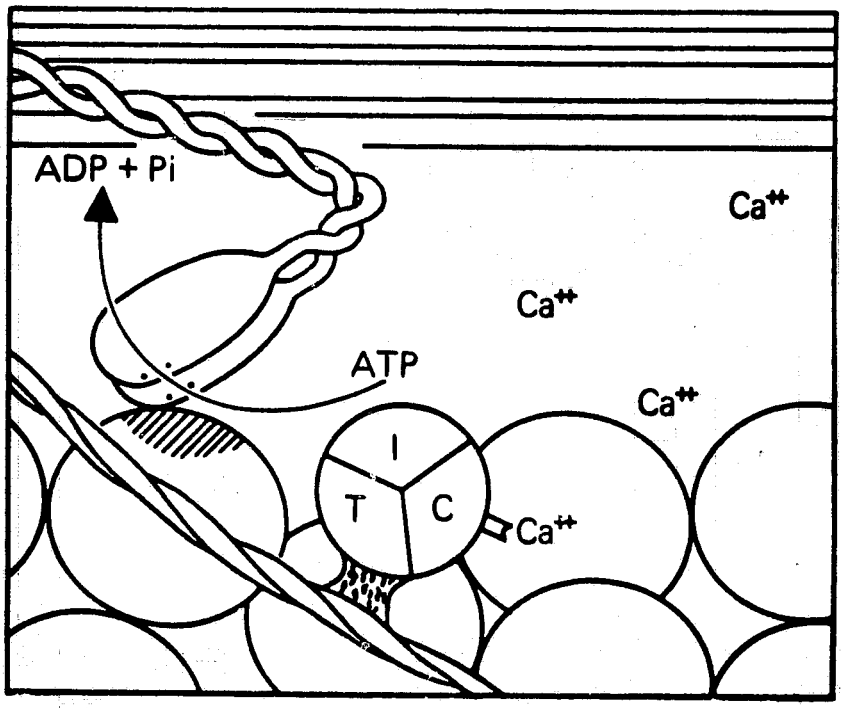


Figure 2: Cross-bridge mechanism of muscle contraction, including possible role of regulatory proteins. (From Ganong,W.F., Review of Medical Physiology)

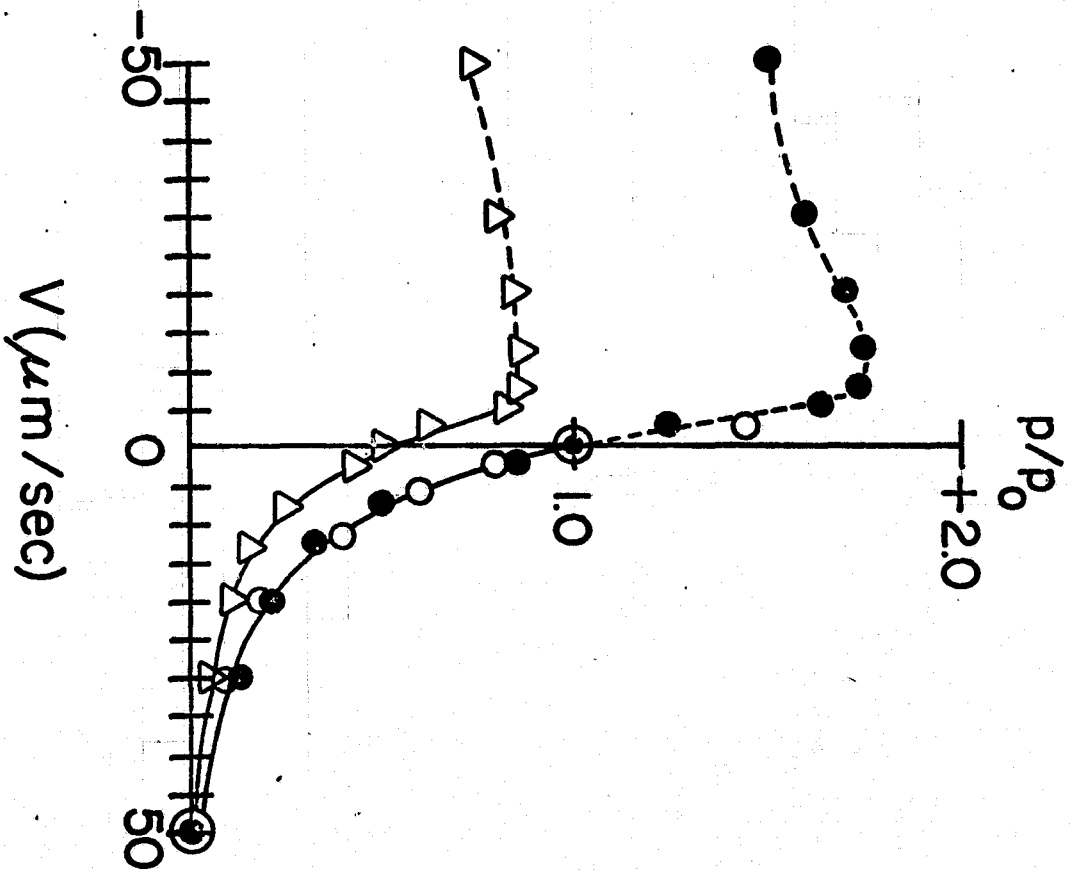


Figure 3: Force-velocity relat ionship for a fast-twitch muscle (rat EDL).
Open circles: experimental data
Closed circles: model simulation with full activation.
Triangles: model simulation with half activation.
Model is two-state cross-bridge.

A MODEL OF THE HUMAN OBSERVER AND DECISION MAKER

by P.H. Wewerinke

National Aerospace Laboratory NLR
The Netherlands

SUMMARY

In this paper a model is described of the human observer and decision maker monitoring a dynamic process. The decision process is described in terms of classical sequential decision theory by considering the hypothesis that an abnormal condition has occurred by means of a generalized likelihood ratio test. For this, a sufficient statistic is provided by the innovation sequence which is the result of the perception and information processing submodel of the human observer. On the basis of only two model parameters the model predicts the decision speed/accuracy trade-off and various attentional characteristics.

A preliminary test of the model for single variable failure detection tasks resulted in a very good fit of the experimental data.
In a formal validation program a variety of multivariable failure detection tasks was investigated. The task variables were the number, the bandwidth and the mutual correlation of display variables and various failure characteristics.
A very good overall agreement between the model and experimental results showed the predictive capability of the model. In addition, the specific effect of almost all task variables was accurately predicted by the model.

INTRODUCTION

With increasing complexity and automation of man-machine systems the human operator's role shifts from controller to supervisor. In the context of transport aircraft operation this is very much the case after the introduction of automatic approach and landing systems and the future microwave landing system (MLS).
The last two decades considerable research effort has been devoted to the study of human control behavior. One result is a number of mathematical tools, of which the state-space, time-domain optimal control model has been shown to provide a general framework adequately describing the human processing of information provided by a dynamic system (Refs 1-4). This can be extended to other cognitive functions involved in monitoring an automatic system, detecting system failures, making decisions, etc. The insight in this higher mental functioning is still rather incomplete although some attempts have been made to investigate and to model failure detection and simple decision making behavior (Refs 5-9).

This paper summarizes the results of a theoretical and experimental

analysis of the human observer and decision maker in multivariable failure detection tasks. In the next section a model of the human decision maker is formulated in terms of multivariable classical sequential decision theory, accounting for the important effect of correlated information. In the subsequent section the model is tested against the results of a single variable task experiment reported in reference 8. Next, a formal model validation experiment is discussed. The latter results are extensively presented in reference 10.

MODEL OF THE HUMAN OBSERVER AND DECISION MAKER

It is assumed that the human perceives information of a linear dynamic system which is described by a Gauss-Markov random sequence. Based on the known dynamics of this system and the perceived information (i.e. noisy observations), the human makes the best estimate of the system state. This is described in standard linear estimation theoretical terms (Kalman filter, Refs 11-12) and is part of the well documented optimal control model (Refs 1-2) but included in figure 1 for the ease of reference.

Now, in the normal mode of operation, the discrepancy between perceived and expected information (the so-called innovation sequence n_k) is a zero mean Gaussian purely random sequence (Ref. 12) with covariance N_k. It is assumed that abnormal system operation, as caused by errors in display instruments, malfunctioning of the system and excessive system disturbance levels (e.g., large windshears in aircraft operation) can be represented by a deterministic process, as such unknown to the human observer but detected on the basis of a non-zero mean innovation sequence whose statistic is sufficient to make decisions (test hypotheses) when the system is completely observable.

In terms of classical sequential decision theory (Refs 13-14) a so-called generalized likelihood ratio test can be formulated. The test amounts to the comparison of the probability of a non-zero mean with the probability of a zero mean innovation sequence assuming that the human operator makes a short-term estimate of the mean of the innovation sequence on the basis of the sample mean of m past observations ($\bar{n}_k$ in figure 1).
It can be derived (Ref. 15) that the effect of each observation at stage k on the (log of the)likelihood ratio is given by

$$\Delta L_k = \tfrac{1}{2}\, \bar{n}_k'\, N_k^{-1} \bar{n}_k \qquad (1)$$

under the assumption that the sample mean $\bar{n}_k$ is constant during m observations. The accumulating effect of each observation on the total (log of the) likelihood ratio is given by the recursive expression

$$L_k = L_{k-1} + \Delta L_k \qquad (2)$$

assuming that the innovation sequence is a white noise sequence (independent samples) which is exact in the normal mode of operation. The number of observations, based on which the decision is made, is chosen such that ΔL_k is, on

the average, not decreasing; in other words, using only confirming evidence to make the decision. Otherwise, the positive semi-definite elements ΔL_k would have an accumulating effect on the likelihood ratio.
When the likelihood ratio L_k (representing the total evidence of abnormal system operation) is equal to, or larger than, a decision threshold T, the decision is made that an abnormal condition has occurred. This decision threshold can conveniently be related to the accepted (or assumed) risk according to (Ref. 14) $T=(1-P_M)/P_F$, with P_M the miss probability (i.e. of no response to an abnormal condition) and P_F the false alarm probability.

Following reference 14 an expression can be derived (Ref. 15) for the average number of samples used to make the decision that the system is operating abnormally, which is, for a given sample rate, uniquely (linearly) related to the average detection time.
This average number of observations K, given the abnormal condition, is given by

$$\bar{K}= \frac{2P_F T\ell nT}{E\{\bar{n}'\bar{N}^{-1}\bar{n}\}} \tag{3}$$

where $\bar{(.)}$ indicates the average over the ensemble and $E\{(.)\}$ is the average over the sequence.
Equation (3) gives a relationship between the average number of observations and the decision error probabilities (P_F and P_M), for a given innovation covariance N and the non-zero mean failure state sequence which is, however, a given task variable. Thus the only human decision model parameters are the short-term average sample size m and the innovation covariances which depend exclusively on the human observation noise covariances (Ref. 15). The model output is the average failure detection time corresponding with given (or assumed) error probabilities.

Previous studies (Ref. 2) support the hypothesis that the observation noise covariance scales with the mean-squared value of the corresponding signal. Thus for display variable j $V_j = P_0 E\{y_j^2\}$, $j= 1,....\ell$. P_0 represents a nominal, "full attention" noise ratio (typically 0.01π). In case a display variable is not looked at (foveally), it is assumed that the corresponding observation noise is infinite, thus neglecting peripheral viewing. Inserting this expression in equation (3) results, after some matrix manipulation (Ref. 15) in a very simple expression for the average "detection time"

$$\bar{K}= \frac{2P_F T\ell nT}{E_t\{\bar{n}_e^2\}} \tag{4a}$$

with

$$\bar{n}_{e_j}^2 = \bar{n}_{r_j}^2 (1-p_{r_j}) \tag{4b}$$

is the effective, relative, estimated sample mean-square; the subscript r indicates the normalization by the (constant) observation noise covariance V_j, p_j is the estimation error covariance of display variable j, and E_t indicates the average over the ensemble, and the sequence, and the display variables.

The expected value of $\tilde{n}_e^2$ over the display variables involves the probability distribution of the human attention to these variables. When an optimal distribution (i.e., yielding a minimal detection time) of the human attention is assumed, the only remaining human decision model parameters are the short-term average sample size m and the overall level of attention P_0.

The general model structure (the only assumptions are that the dynamic system is linear and that abnormal conditions can be represented by a deterministic process) accounts for the effect of a variety of task variables such as the number and bandwidth of display variables, the correlation among them and the perceived failure characteristics. These variables are included in the selected task configurations investigated in the validation program discussed in the next section.

MODEL VALIDATION

In order to test the validity of the model the results of two experimental programs were considered and compared with the model predictions. The first experiment is reported in reference 8 in which observers were required to detect a change in the mean of a stationary stochastic process. The experimental results were used for a preliminary validation of the human decision model and to "calibrate" the model with respect to the short-term average sample size m. Next, a formal model validation program is described which is reported in reference 15.

Preliminary model validation

In the first experiment the (two) subjects were instructed to detect, as soon as possible, the occurrence of a non-zero mean component while observing a second order, zero mean process. Both step and ramp failures with four possible amplitudes were introduced yielding 8 experimental conditions. Model results of the average failure detection time were obtained assuming the decision error probabilities of 0.05 which were also obtained in the experiment and on the basis of a typical overall level of attention P_0 of 0.01π. The remaining model parameter m was selected so as to obtain the best overall match with the experimental results. As shown in figure 2 the resulting value of m of four seconds yields an excellent fit to the experimental data. This can be expressed in the linear correlation coefficient between model and experimental results of 0.99 and in the ratio of model detection times and corresponding experimental values (t_m/t_e). This ratio was on the average, 0.98, with a standard deviation of 0.09. Thus the constant value of m of 4 seconds yielding a good fit for all the experimental conditions seems a human operator-related parameter. This value of four seconds, which ties in very well with the short-term memory span typically ranging up to five seconds for visual stimuli (Ref. 16), will be assumed and kept constant in following validation experiment.

Formal model validation

A variety of tasks was specified so as to present the most crucial combination of the task variables considered: number, bandwidth and mutual correlation of the displays and failure characteristics. Two failure types were investigated either appearing on one display (display failure, F_i) or on two correlated displays (system failure, F_s). Also the effect of prior knowledge about the failure type was investigated.
Up to four-display tasks were considered consisting of two separate (independent) identical processes. Each process could be observed via two displays: a relatively high bandwidth variable y1 (second order process with a break-frequency of 1.2 rad/s) was additionally filtered (first order filter with a time constant of four seconds). The output of this filter which is correlated with y1 ($r = 0.5$) was displayed as y2. This process was duplicated resulting in a four-display process (y1 to y4), of which the displays were two by two correlated.
The resulting 8 configurations and display situation are summarized in table 1. The configurations were investigated for two failure rates (a ramp with a slope of 0.1 standard deviation of the display position per second and a slope of 0.2 σ_{y_i}/s) resulting in 16 experimental conditions. A system failure appeared as a ramp on y1 which was additionally filtered and subsequently displayed. For a detailed presentation of the foregoing tasks and the model and experimental results (of three subjects, being general aviation pilots; twelve replications per condition), the reader is referred to reference 15. In this paper only the principal results are summarized aimed at a test of the human decision model.

Failure detection times

Model predictions of the (ensemble) mean failure detection times for all the 16 failure detection tasks were obtained on the basis of the two constant model parameters: the overall level of attention $P_0 = 0.01\pi$ and the short-term average sample size $m = 4$ seconds. Based on the previous results a value of the false alarm probability of 0.05 was assumed. Additional model assumptions and procedural details are discussed in reference 15.
The results of two subjects achieving an overall false alarm probability P_F of 0.05 could be compared directly with the model predictions. The result is summarized in figure 3. The linear correlation coefficient between the model predictions and the experimental mean failure detection times is 0.86 which reflects a very good overall predictive capability of the human decision model. The ratio of the model and experimental failure detection time t_m/t_e is an other measure for the agreement between the model and experimental results. On the average (over all 16 tasks) this ratio is 0.98. The standard deviation is 0.12 which is comparable with the reliability of the experimental estimated means.
This reliability of the data is included in figure 4 showing the model and experimental failure detection times per configuration. Apart from the mean value, also the standard deviation of the mean value estimate ($\sigma/\sqrt{N}$, with σ the standard deviation of the raw data) is given. For almost all configurations

the model predictions agree rather well with the experimental failure detection times.
For one configuration (Conf. 4) the model predictions clearly disagree with the experimental results. A very plausible explanation of this discrepancy (which was confirmed during the debriefing of the subjects) was that the subjects did not realize (use) the system failure dynamics but assumed that the system failure appeared simultaneously on both (correlated) displays. The model results based on this assumption are also shown in figure 3 and indicated with model refinement. In that case, the linear correlation coefficient is 0.95; the mean ratio t_m/t_e is 1.01 with a standard deviation of 0.09.

The agreement between the model predictions and experimental results with respect to the specific task variables is summarized in table 2. The configurations involved in the pair-wise comparison between the configurations which differ with respect to the specific task variable (only) are indicated.
Comparing the measured and model results shows that the effect of display bandwidth, of additional (correlated) displays and of the failure rate is excellently predicted by the model. The predicted interference between uncorrelated displays (because of the human attention sharing involved) is larger than obtained experimentally. The model predictions are based on a constant level of attention. However, the physiological (heart rate) measures obtained during the experiment suggest a small, but statistically significant, increase in attention with an increase in displays which can easily explain the small difference in interference. The effect of the failure type is discussed before. It can be seen that the model refinement and the experimental results agree closely.
The experimental results of the third subject reflect a distinctly different decision strategy. He made no false alarms and his failure detection times were, on the average, 40 % higher. Yet his results correlated well with the model predictions ($r = 0.80$). Model results based on a very low value of P_F yield an arbitrary good overall correspondance with the measured failure detection times (detection times increase monotonically with decreasing P_F). However, virtually the same linear correlation coefficient is obtained. As discussed in reference 15 this signifies that the predicted effect of the various configurations on the failure detection time match the results of this subject with an accuracy of about 12 % (the standard deviation of t_m/t_e is 0.12).

Scanning behavior

Various attentional characteristics can be derived from the foregoing model of the human observer and compared with eye scanning data obtained in the experimental program.
Combining eqs (1), (3) and (4), it can easily be seen that the (ensemble) average effect on the likelihood ratio of one observation of display variable j is given by

$$\overline{\Delta L}_j \doteq \tfrac{1}{2} \overline{\dot{n}^2}_{e_j} \, .. \overline{\delta}_j \qquad (5a)$$

where $\bar{\delta}_j$ indicates the probability of attending to j. In the normal mode of operation ($\bar{n}$= 0) an alternative expression can be obtained (Ref. 15)

$$\overline{\Delta L}_j \underset{\text{mode}}{\overset{\text{normal}}{=}} \tfrac{1}{2} \Delta p_{r_j} \cdot \bar{\delta}_j \qquad (5b)$$

where Δp_{r_j} is the reduction in the estimation error covariance due to observing variable j.

Equations (4) and (5) show that the optimal allocation of attention among the displayed variables (i.e. yielding the maximum $\overline{\Delta L}$ and thus, the minimal detection time) is obtained for the maximum expected value of $\bar{n}_e^2$. The model predicts (Ref. 15) that the optimal fraction of attention to the high bandwidth display(s) is varying between 0.4 and 0.7 (the total attention to the high and low bandwidth display(s) is 1.0) somewhat depending on task specifics and failure characteristics. As the failure detection time is relatively insensitive to the division of attention in this region, a relatively constant fraction of attention to the high bandwidth display(s) for all configurations is predicted by the model, say between 0.5 and 0.6 (enhanced by the randomized block design and the corresponding transfer of training). This agrees very well with the experimental dwell fractions. The average dwell fraction on the upper display was not depending on the configurations and varied between 0.5 and 0.55.

The optimal allocation of attention in the time domain - thus the optimal scanning strategy for a given task - is described by eq. (5). Various model predictions can be derived from the attention allocation model and compared with the corresponding eye scanning measures. This is discussed in reference 15. For illustrative purposes, consider the normal mode of system operation, for which situation the effect of observing is described by eq. (5b). The model predicts that there is a scanning preference for high bandwidth display variables as the estimation error covariance Δp increases with display frequency, for a given amount of time. Furthermore, it follows from eq. (5b) that there is a scanning preference for correlated display variables, because observing variable j yields an additional reduction in Δp_i if variable j is correlated with i which results in a maximum total ΔL and a minimum (average) failure detection time.

These model predictions can be compared with the experimental eye scanning data in terms of display link values. The experimental results agreed well with these qualitative predictions. The fraction of links between the correlated displays was about two times the fraction of links between the uncorrelated displays (0.64 versus 0.36; in the failure mode this division is 0.70 versus 0.30 which increase is predicted on the basis of eq. (5a) as discussed in reference 15). Furthermore, the fraction of scans (number of observations) towards the high bandwidth displays was two times the fraction of scans towards the low bandwidth displays.

The foregoing analysis illustrates the predictive capability of the attention allocation model which may be a powerful tool in the study of human information processing tasks and display design problems.

CONCLUDING REMARKS

In this paper a model of the human observer and decision maker is summarized. The model consists of two parts. A submodel of the human observer is formulated in linear estimation theoretical terms including the perception of the displayed information of a linear(ized) process and the information processing stage which is described by a Kalman filter. The resulting innovation sequence provides a sufficient statistic for the decision process. In terms of classical sequential decision theory the hypothesis is considered that an abnormal condition has occurred by means of a likelihood ratio test. An abnormal condition is represented by a deterministic process which has to be detected on the basis of noisy observations of a normally zero-mean stochastic process. On the basis of only two model parameters (short-term average sample size m and the overall level of attention P_0) the model predicts the (ensemble) mean failure detection time and various attentional characteristics.

A preliminary test of the model for single variable failure detection tasks resulted in a very good fit of the experimental results for a constant value of m of four seconds.
This constant value for m and a typical value for the overall level of attention were used to predict the mean failure detection times and scanning characteristics of a variety of multivariable failure detection tasks which were investigated in the formal validation program. The task variables were the number, the bandwidth and the mutual correlation of display variables and various failure characteristics.
A very good overall agreement between the model and experimental results both in terms of failure detection times and eye scanning measures showed the predictive capability of the model. In addition, the specific effect of almost all task variables was accurately predicted by the model.

REFERENCES

1. Kleinman, D.L. and Baron, S.: Manned Vehicle System Analysis by means of Modern Control Theory. NASA CR-1753, June 1971.

2. Baron, S. and Levison, W.H.: Display Analysis with the Optimal Control Model of the Human Operator. Human Factors, 1977, 19(5).

3. Wewerinke, P.H.: Performance and Workload Analysis of In-Flight Helicopter Missions. Paper presented at the 13th Annual Conference on Manual Control, MIT, June 1977 (also NLR MP 77013 U).

4. Wewerinke, P.H.: The Effect of Visual Information on the Manual Approach and Landing. Paper presented at the 16th Annual Conference on Manual Control, MIT, May 1980 (also NLR MP 80019 U).

5. Phatak, A.V. and Kleinman, D.L.: Current Status of Models for the Human Operator as a Controller and Decision Maker in Manned Aerospace Systems. AGARD symp. on Automation in Manned Aerospace Systems. WPAFB, Dayton, USA, 1972.

6. Levison, W.H. and Tanner, R.B.: A Control-Theory Model for Human Decision Making. NASA CR-1953, December 1971.

7. Wewerinke, P.H.: Human Monitoring and Control Behavior - Models and Experiments. NLR TR 77010 U, October 1976.

8. Gai, E.G. and Curry, R.E.: A Model of the human Observer in Failure Detection Tasks. IEEE Trans. on Systems, Man and Cybernetics. Vol. SMC-6 no. 2, February 1976.

9. Gai, E.G. and Curry, R.E.: Failure Detection by Pilots during Automatic Landing: Models and Experiments. J. Aircraft, Vol. 14, no. 2, February 1977.

10. v.d. Graaff, R.C.: An Experimental Analysis of Human Monitoring Behavior in Multivariable Failure Detection Tasks. NLR TR 81063 U, June 1981.

11. Bryson, A.E. and Ho, Y.C.: Applied Optimal Control. Blaisdell Publishing Company, 1969.

12. Mehra, R.K. and Peshon, J.: An Innovation Approach to Fault Detection and Diagnosis in Dynamic systems. Automatica, Vol. 7, 1971.

13. Wald, A.: Sequential Analysis, J. Wiley, New York, 1947.

14. Sage, A.P. and Melsa, J.L.: Estimation Theory with Applications to Communications and Control. McGraw-Hill, 1971.

15. Wewerinke, P.H.: A Model of the Human Decision Maker Observing a Dynamic System. NLR TR 81062 U, March 1981.

16. Sheridan, T.B. and Ferrell, W.R.: Man-Machine Systems: Information, Control and Decision Models of Human Performance. MIT Press, Cambridge, Ma. 1974.

Table 1a
Task configurations

CONF.	DISPLAY	FAILURE
1	y1	F_1
2	y2	F_2
3	y1, y2	F_1
4		F_s
5		F_1, or F_2, or F_s
6	y1, y3	F_1 or F_3
7	y1, y2, y3, y4	
8		F_i, or F_{s_j}

Table 1b
Display situation

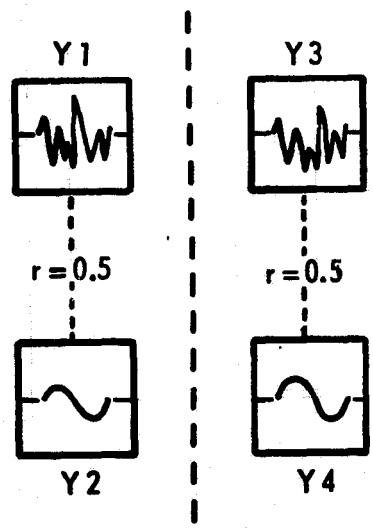
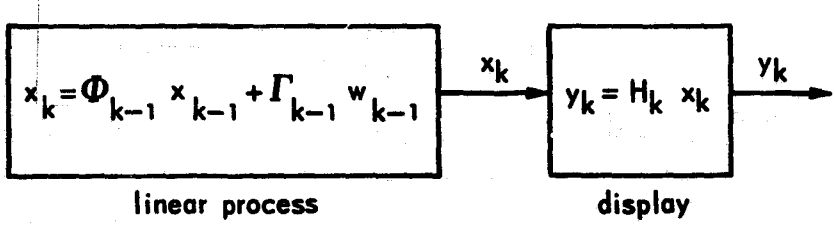


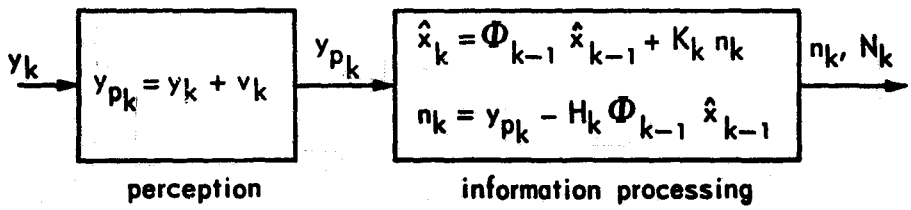
Table 2
Specific effect of task variables on model and experimental failure detection times

EFFECT OF	CONFIGURATIONS	RATIO t_i/t_j			
		measured		model	
		m	s	m	s
bandwidth (low/high)	1, 2, 5, 8	1.01	0.12	1.00	0.08
additional display info.	1, 3, 6, 7	0.76	0.06	0.77	0.07
inter-ference	1, 3, 5, 6, 7, 8	1.11	0.05	1.24	0.05
failure type (system/ display)	3, 4, 5, 8	1.43	0.33	1.24 (1.37)	0.06 (0.21)
failure rate (high/low)	all	0.67	0.05	0.69	0.07

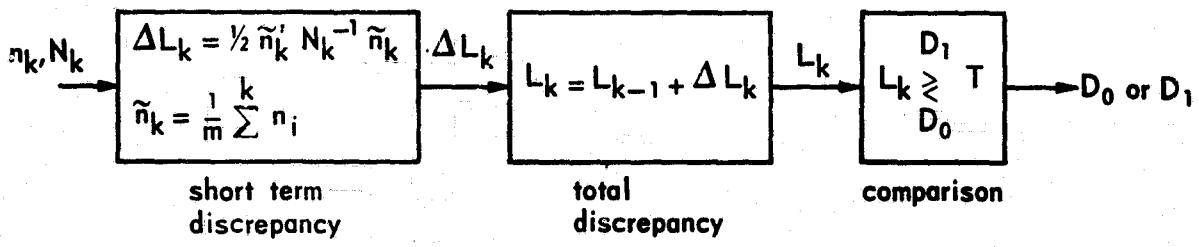
(.) model refinement



a) dynamic system



b) human observer



c) human decision maker

Fig. 1 Model of the human decision maker observing a dynamic process

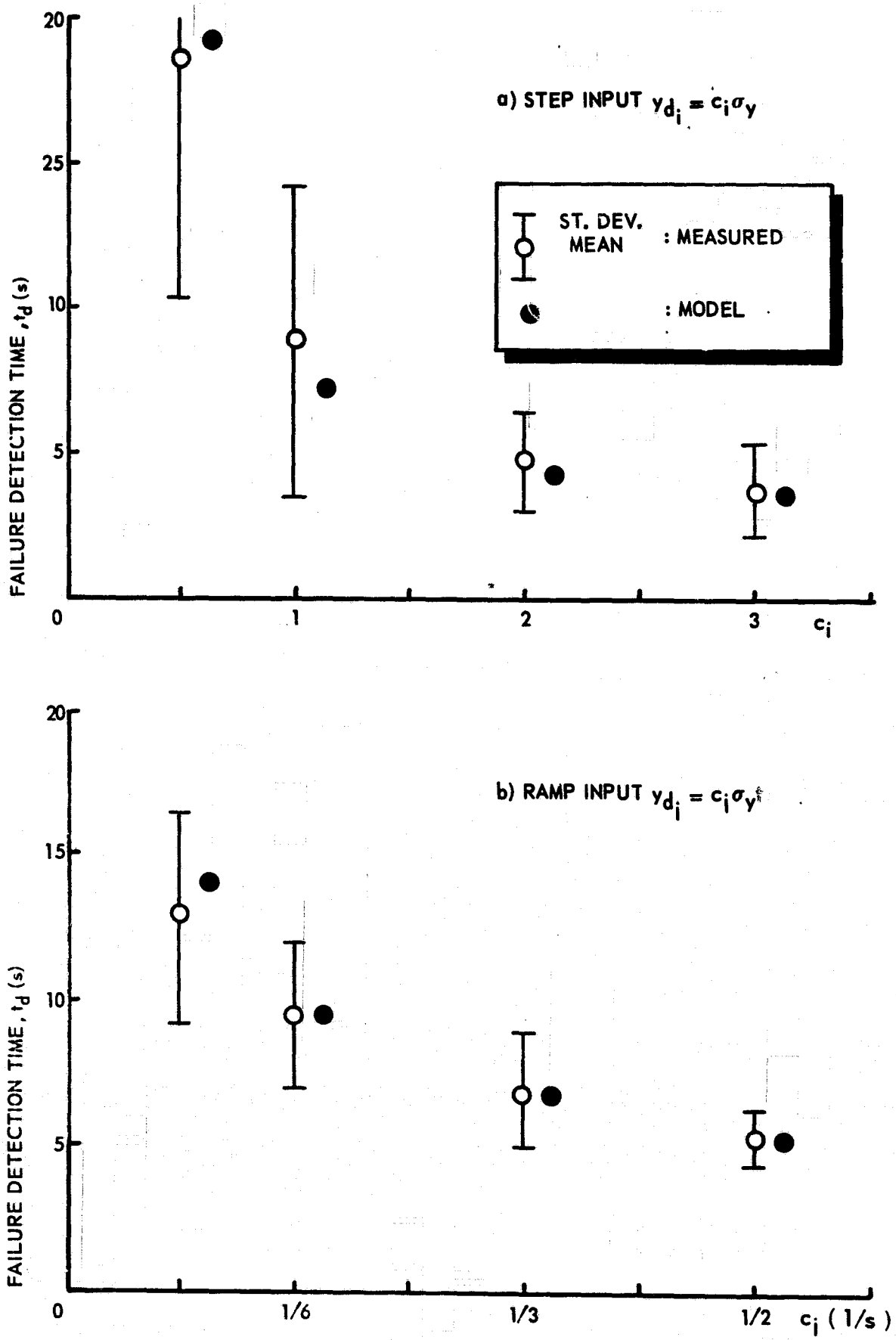


Fig. 2 Experimental and model failure detection times for various step and ramp inputs

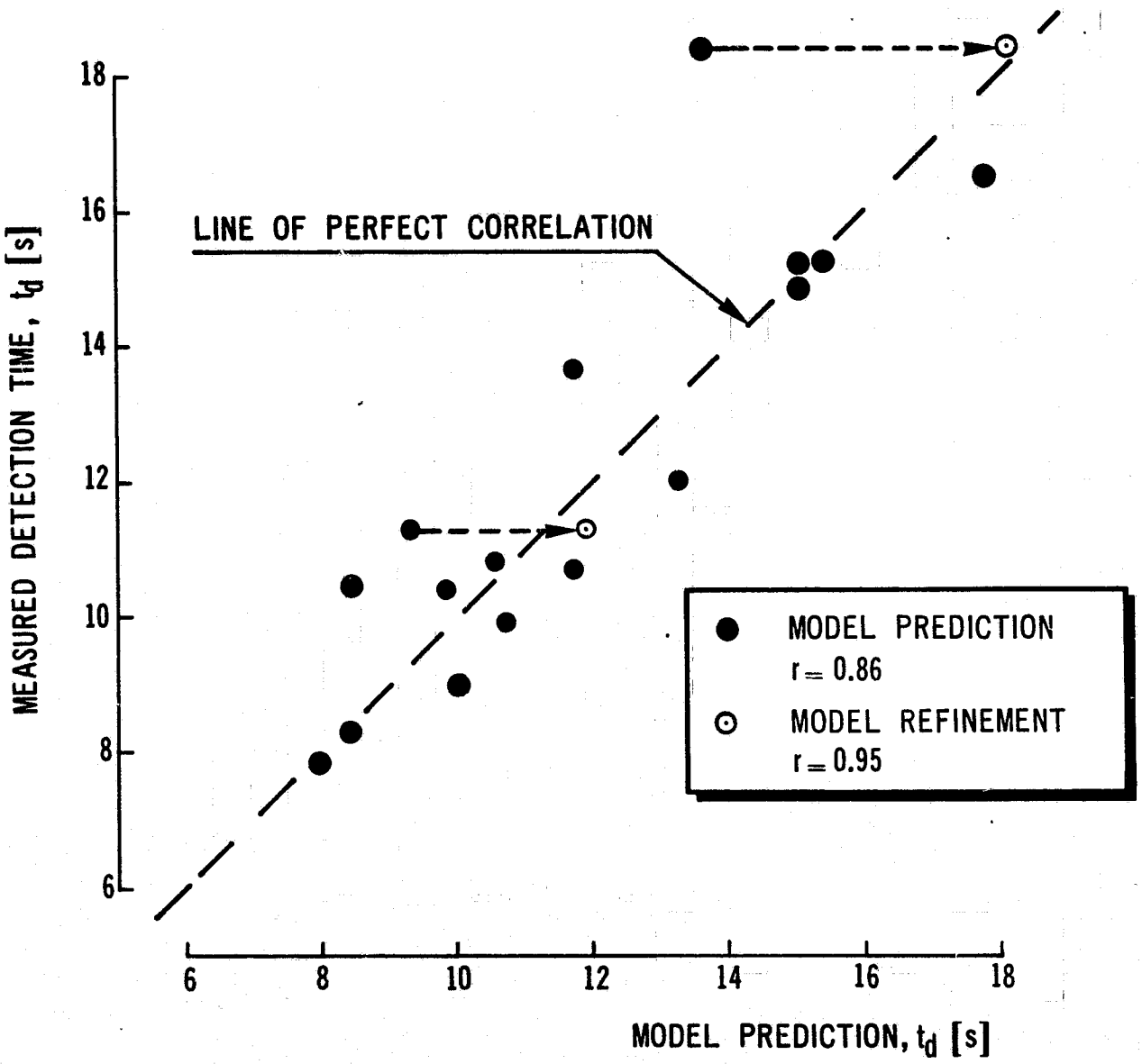


Fig. 3 Overall comparison of model predictions and experimental results of two subjects

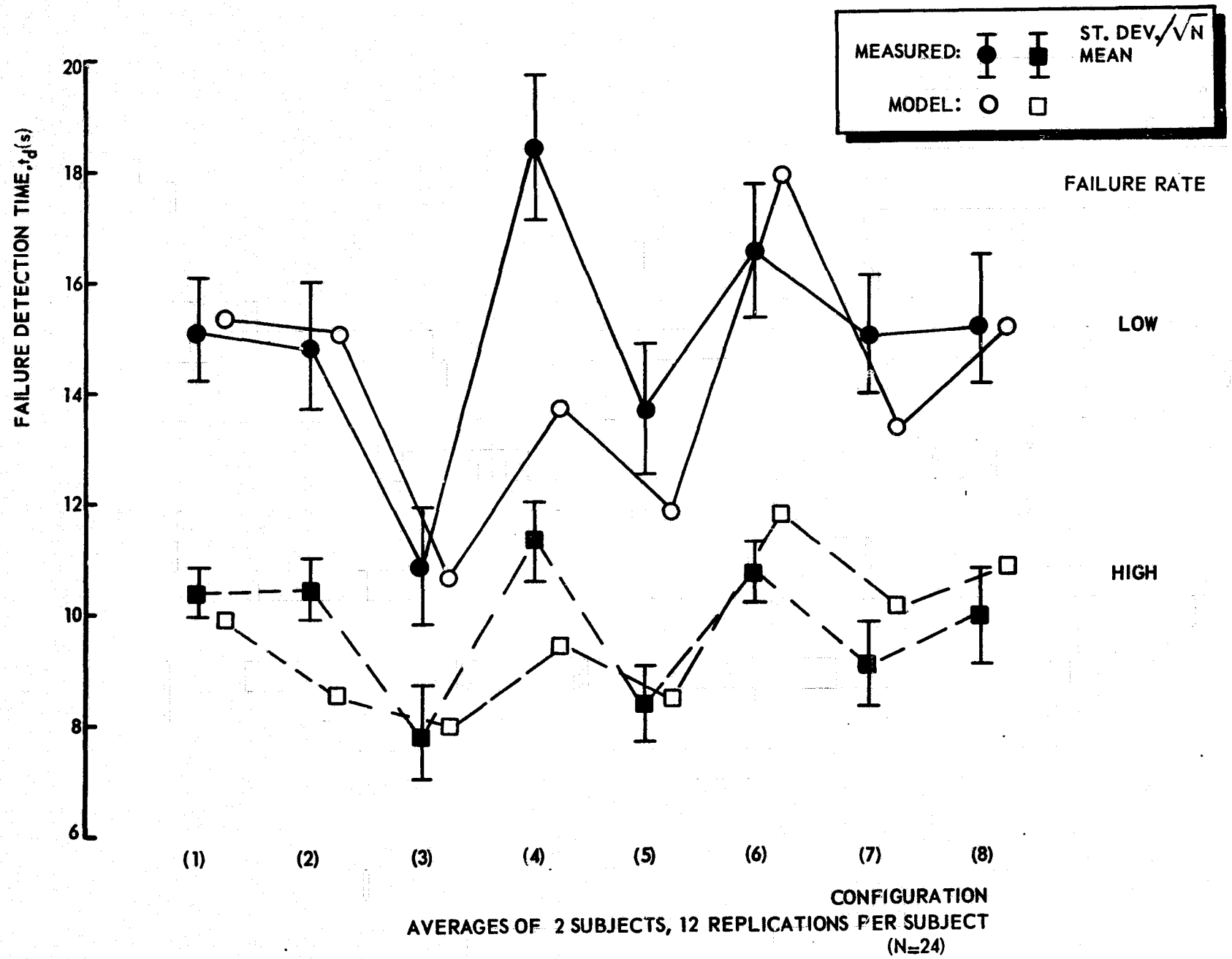


Fig. 4 Model and experimental failure detection times for all configurations

D50

HUMAN SUPERVISION AND MICROPROCESSOR CONTROL OF AN OPTICAL TRACKING SYSTEM

By William J. Bigley and John D. Vandenberg

Lockheed Electronics Company

SUMMARY

Gunners using small calibre anti-aircraft systems have not been able to track high-speed air targets effectively. Substantial improvement in the accuracy of surface fire against attacking aircraft has been realized through the design of a director-type weapon control system. This system concept frees the gunner to exercise a supervisory/monitoring role while the computer takes over continuous target tracking. This change capitalizes on a key consideration of human factors engineering while increasing system accuracy. The advanced system design, which uses distributed microprocessor control, is discussed at the block diagram level and is contrasted with the previous implementation.

INTRODUCTION

The allocation of system functions between man and machine is one of the key elements of human factors engineering. Typically, this allocation is determined by various technical specialists working closely together as an interdisciplinary team. One of the problems faced by system designers is that technical, cost, and operational constraints acting alone or in combination may override human factors considerations thereby jeopardizing the effectiveness of a system. This situation has plagued the designers of closed-loop target tracking systems which, for one reason or another, must use the human operator as a key element in the tracking loop.

Very early gunfire control systems gave the human operator no tracking assistance whatever. The operator was required to estimate all lead and superelevation angles and, in smaller systems, move the weapon as well. With the advent of power-driven mounts, the operator was no longer required to move the sighting mechanism, but, the difficult task of predicting future target position remained. Although systems designed to estimate future target position automatically contained inherent instabilities, performance improved over earlier systems. On an absolute basis, however, the accuracy of the gunfire delivered against fast-moving targets was still relatively poor. Increasing refinements in control system technology incorporating the use of human operators continued to yield improvements (see references 1 through 6), but manned systems were still less accurate than unmanned ones.

As an alternative to providing the human controller with various forms of assistance, for non-maneuvering targets, the principle of decoupling the gunner from the loop was advanced by Lockheed Electronics Company (LEC), where planning sessions dealt with the development of an improved small calibre weapon control system using optical tracking techniques. But a completely automated system did not meet program requirements, so another alternative was suggested: "Since we must have an operator in the tracking system, can he be removed immediately after target acquisition and retained as a performance monitor to supply minor corrections if the occasion demands them?" If applied successfully, such a solution would enable the system designers to free the operator from repetitive tasks while improving the performance of the system. The operator would initiate the track and then be removed quickly from the primary loop when the automated system could maintain the track. The operator could retain an overview sufficient to permit the simpler task of inserting corrections when needed. If the concept could be applied successfully, the goal of many man-machine system designers would be achieved: The operator would supervise system activity and not be burdened with the constant and demanding task of processing system information. Through such reallocation of system functions between man and machine, a low-cost, highly accurate manned fire control system could be designed. The usual role of the operator as an integral element in the control loop would be changed by deemphasizing track loop dependence on the non-linear human transfer function.

After a number of design studies, computer simulations, and laboratory implementations, a design emerged that worked so well in practice that U.S. Patent No. 4,004,729 (reference 7) was awarded. The system performs as well as a fully automatic system and has been implemented in an existing U.S. Army anti-aircraft weapon. A shipboard version is currently in production.

SYMBOLS

E_e	Voltage, tracking error
E_g	Voltage, gun rate loop command
E_R	Voltage, rate gyro
E_s	Voltage, sight rate gyro loop command
V_s	Voltage, sight rate aid
V_g	Voltage, gun rate aid
θ_g	Gun line position angle, with respect to space
θ_s	Sight line position angle, with respect to space
$\theta_{s/g}$	Sight line position angle with respect to gun line ($\theta_s-\theta_g$)

θ_A	Radar line position angle with respect to space
θ_T	Target position angle with respect to space
ϵ	Target tracking position error with respect to space: $(\theta_T-\theta_S)$ in optical track; $(\theta_T-\theta_A)$ in radar track
CX	Synchro transmitter
λ	Lead angle
G_1	Dynamics of director (Laplace Transform)
G_2	Dynamics of plant (Laplace Transform)
K_1, K_2, K_3	Magnitude of gain
S	Laplace Transform ($j\omega$)

MANNED WEAPON CONTROL TECHNOLOGY

Figure 1 is a stylized description of the essential elements of both prior state of the art gunnery and the principles set forth in Patent No. 4,004,729.

System operational concepts are reviewed to permit distinctions to be made between the new system design and earlier designs.

BASIC SYSTEM CONFIGURATION

Many small calibre (20mm to 40mm) gun mounts are manned in order to avoid the higher cost of a remote, full-track gun fire control system (GFCS). By using the operator to perform the target position sensing function in acquiring and tracking the target, an angle tracking radar or electro-optical system is not required, thus significantly reducing system cost. In this case, target range data only is supplied by the radar system or by a laser system.

The present state-of-the-art GFCS configuration consists of five major subsystems:

1. Gun/Turret servos, elevation and azimuth.

2. Optical gunsight servos, elevation and azimuth.

3. Radar Antenna Pedestal servos, elevation and azimuth.

4. Microcomputer, digital.

5. Operator/Stick.

Although the subsystems and their axes can be combined in many geometrical arrangements, a configuration most commonly found is shown in figure 1. The optical gunsight, radar antenna, microcomputer, and operator/stick subsystems are all mounted on the turret azimuth platform along with the gun elevation drive mechanism.

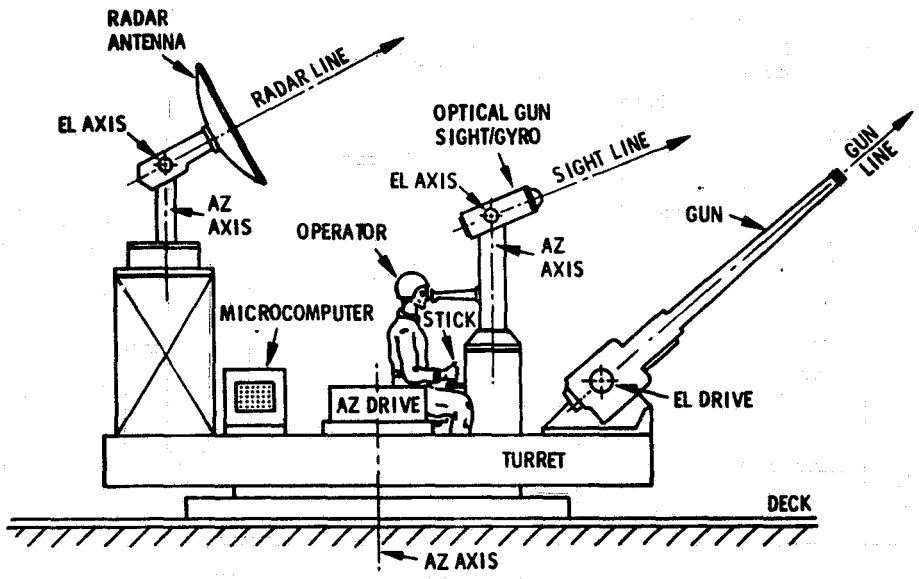


Figure 1. Manned Weapon Control System - Basic Configuration

In the configuration selected for discussion, notice that all three elevation axes are mounted parallel to the turret deck. So, for example, if the gun and radar elevation position servos are slaved to the optical sight position (with respect to the turret deck), the gun line and the radar line would follow and be parallel to the sight line in space (ignoring gun lead angles for the moment). Notice further that both the radar and sight azimuth axes are mounted on top of and parallel to the turret azimuth axis. Therefore, the radar and sight lines in azimuth are not required to be parallel to the gun line in azimuth. (The gun line is actually the turret azimuth position.) So, there is a relative motion problem to contend with in azimuth. Both the sight and the radar azimuth servos are positioned relative to the gun azimuth servo. For large angles of travel, the turret azimuth drive transports both the sight and the radar azimuth servos. In other words, the sight azimuth servo and the radar azimuth servo go along on the turret for a "free ride". It is important to visualize the relative motion relationships among all six axes before trying to understand how the gun fire control system physically functions.

There are many ways in which the five subsystems can be combined and controlled to form a weapon control system. In final analysis, however, the GFCS usually uses the computer and operator/stick subsystems to coordinate and provide hierarchial feedback control of the gun, sight, and radar subsystems in three major modes of operation: (1) target acquisition, (2) target tracking, and (3) ballistics solution and generation of gun lead angles. The scope of this paper is confined to examining man-machine interface configurations in the target tracking mode of operation.

Before presenting a description of the new director-type Sharpshooter system, a brief explanation of the earlier disturbed-line-of-sight system will demonstrate how the director-type control significantly improves the relationship between the human visual/motor system and the machine system.

PRIOR-ART: THE DISTURBED-LINE-OF-SIGHT SYSTEM

Most man-in-the-loop weapon control systems developed in the past can be characterized as Disturbed-Line-of-Sight systems (DLOS). In a DLOS system, the human operator plays a key role by becoming a cascaded component in the target tracking feedback loop as shown in figure 2.

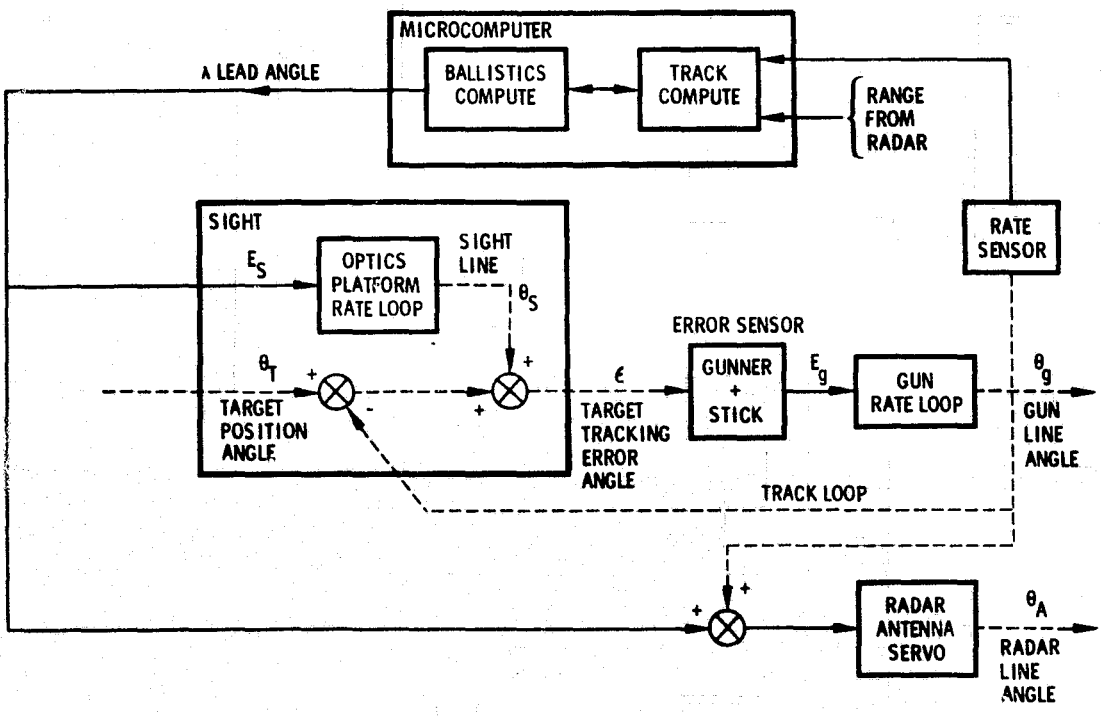


Figure 2. Disturbed Line-of-Sight Track Loop - Conceptual Model

In order to explain how the DLOS system works, consider the case where an airplane (target) is flying at a constant altitude in a circle, at the center of which is the gun mount. Elevation angles of sight, radar, and gun lines are equal and constant. With lead angle λ in figure 2 equal to zero, the gunner looks through the sight and observes the azimuth tracking error $\epsilon=(\theta_T-\theta_g)$ by comparing the position of the target with the center of the gunsight reticle. By controlling the stick voltage outut E_g, the gunner commands the speed of the gun in azimuth to make the speed of the gun line $\dot{\theta}_g$ and the sight line $\dot{\theta}_s$ equal to the speed of the target $\dot{\theta}_T$ when the center of the reticle is as close to the target as the human visual/motor system can make it.

At this point, the computer completes the ballistics solution and starts to inject the azimuth gun lead angle order λ into the sight optical platform as shown in figure 2. The gunner sees the reticle moving off and trailing the target. Consequently, in an effort to get the center of the reticle back on to the target, the gunner adjusts the stick to move the turret faster in the opposite direction, which now moves the azimuth gun line ahead of the target. In other words, the computer is disturbing the sight line so it lags the target, and the gunner is countering the disturbance by commanding the gun line (and sight) to lead the target. This "dispute" over system control between the human operator and the computer continues until a steady state is reached (for this scenario only) where the gunner has the reticle back on the target and the computer has the gun line θ_g leading the sight line θ_s by an angle λ.

The human error sensor processes all of the target track data and the lead angle data. It can be seen that the whole system quickly degrades and approaches instability when the operator has to track a fast maneuvering target.

DIRECTOR TYPE TRACKING SYSTEM

Control Technique: Velocity-Feed-Forward

The design of the director type system depends heavily upon a control engineering technique called velocity-feed-forward. Sometimes this concept is briefly referred to as "rate-aid". The principle of "rate-aid" is illustrated generically in the block diagram in figure 3.

Given a position feedback loop, which is exactly what a tracking loop is, the objective is to reduce the track position error ϵ to an acceptable level. This is usually done by raising the open loop gain (K_1K_2), as long as loop stability is maintained. Velocity-feed-forward makes this possible because it is a function (K_3S in figure 3) that is not inside the loop and therefore does not tend to degrade loop stability.

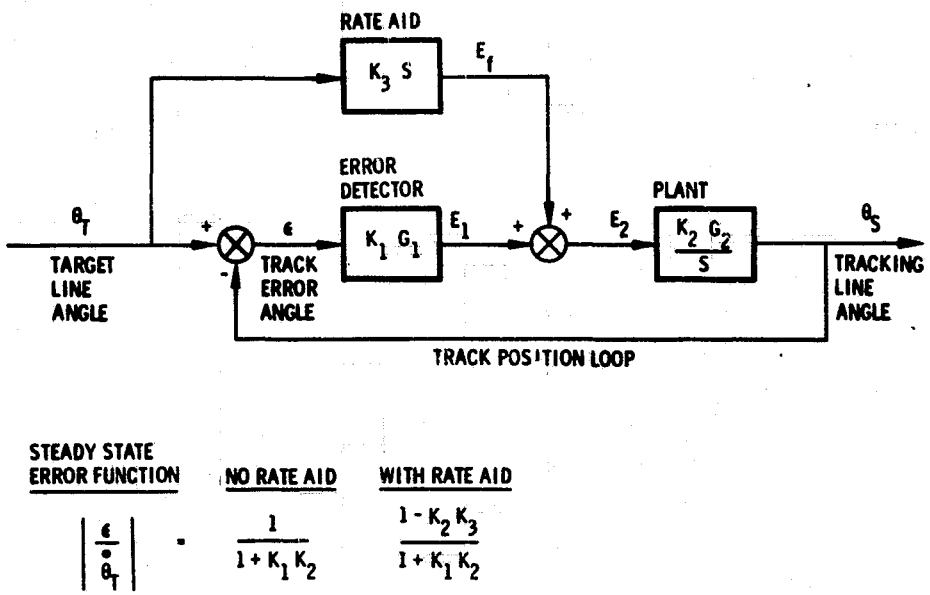


Figure 3. Velocity-Feed-Forword Principle of Dynamic Error Reduction

If the target velocity $\dot{\theta}_T$ is once again constant, say 0.5 radian/second, the plant input voltage E_2 must be 5 Vdc in order to make $\dot{\theta}_S$ = 1 radian/second. Without velocity-feed-forward, the track error ϵ must be a certain value in order to make E_2 = 5 Vdc. The value of ϵ is determined by the transfer function shown in figure 3 for no rate aid.

With the rate aid (K_3S) function scaled properly, E_2 can be set to 5 Vdc by letting the rate aid output voltage E_f provide the 5 Vdc. Now the error detector voltage E_1 and the track error ϵ can go to zero. This relationship is shown in the figure 3 transfer function with rate aid.

Sharpshooter: The Manned Director-Type Tracking System

Referring once more to the basic weapon system shown in figure 1, the director type tracking system uses the relative motion relationship between the sight and radar azimuth axes and the turret axis and the velocity-feed-forward control technique to track and generate gun lead angles without either the operator/stick or the computer subsystems being inside the loop. A conceptual block diagram for the director tracking system is shown in figure 4. (To simplify the diagram, details such as D/A converters have been eliminated.)

In the figure 4 system, the radar subsystem provides range data to the computer for ballistics and is slaved to the sight so that θ_A=θ_S at all times. Note also in figure 4 that a rate gyro has been added as feedback around the sight optics platform. The sight line is inertially stabilized in space, and when the turret azimuth axis θ_g moves under the sight azimuth axis (see figure 1), the sight line θ_S will not move with θ_g.

The major difference to note between the systems shown in figure 4 and figure 3 is that the track error sensor (in this case the operator) is not in cascade between the sight and the gun loops. As we shall see, the operator/stick subsystem has been ideally eliminated from the loop, even though the operator is still observing ϵ, the tracking error, as seen through the sight reticle.

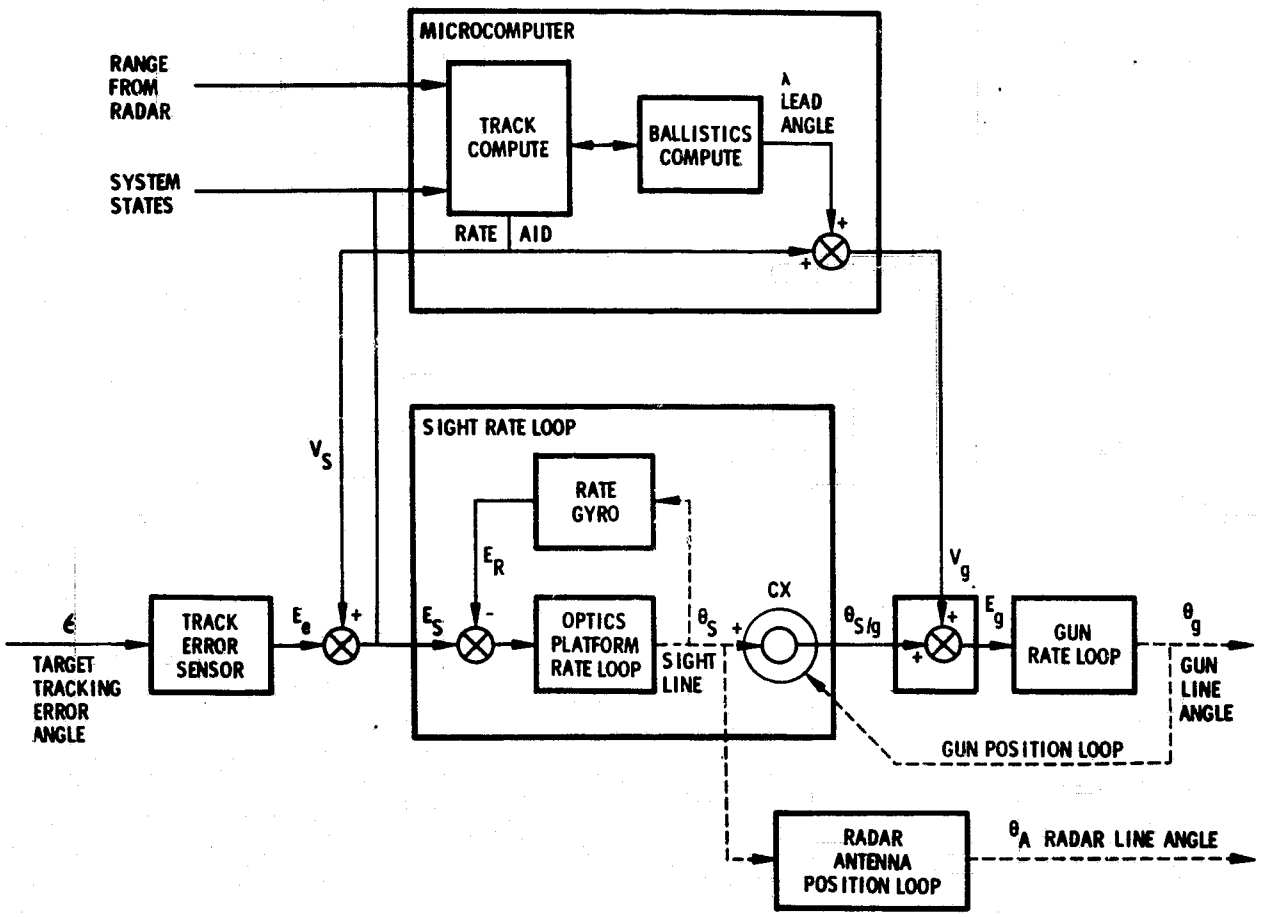


Figure 4. Sharpshooter Director-Type Track Loop - Conceptual Model

For purposes of discussion, consider again the target at constant velocity and range in order to simplify the explanation of how the manned director tracking system functions physically. Diagrams of the tracking sequence are presented in figure 5.

Assume the target is flying at a fixed altitude in a circular path (fixed range) at a constant angular velocity of 0.5 radians/second. At the center of the circle is the weapon system in figure 1 with the fire control system in figure 4. When the operator initiates tracking, the track error sensor (operator/stick) in figure 4 senses error ϵ (difference between target and reticle functions) and enters a correction. This causes a stick output voltage to be applied to the gyro stabilized sight loop (E_s). The

sight line (θ_s) moves ahead of the gun line (θ_g), causing this angular difference ($\theta_{s/g}$) in the sight synchro transmitter (CX) to apply a voltage E_g to the gun rate loop. In order to make the gun line (turret) velocity θ_g reach 0.5 radian/second, assume E_g = +5 Vdc is required. When the sight line θ_s has moved ahead of the gun line θ_g such that the sight-with-respect-to-gun angle $\theta_{s/g}$ is at the value to produce +5 Vdc at E_s, then the relative motion between the sight azimuth and turret azimuth axes stops. At this point, all three axes are rotating in space at 0.5 radian/second; the sight reticle is (depending upon operator dexterity) either on the target or lagging slightly at an angle ϵ; the relative position between the three lines is as shown in figure 5(a); and the operator is providing 5 Vdc at E_s.

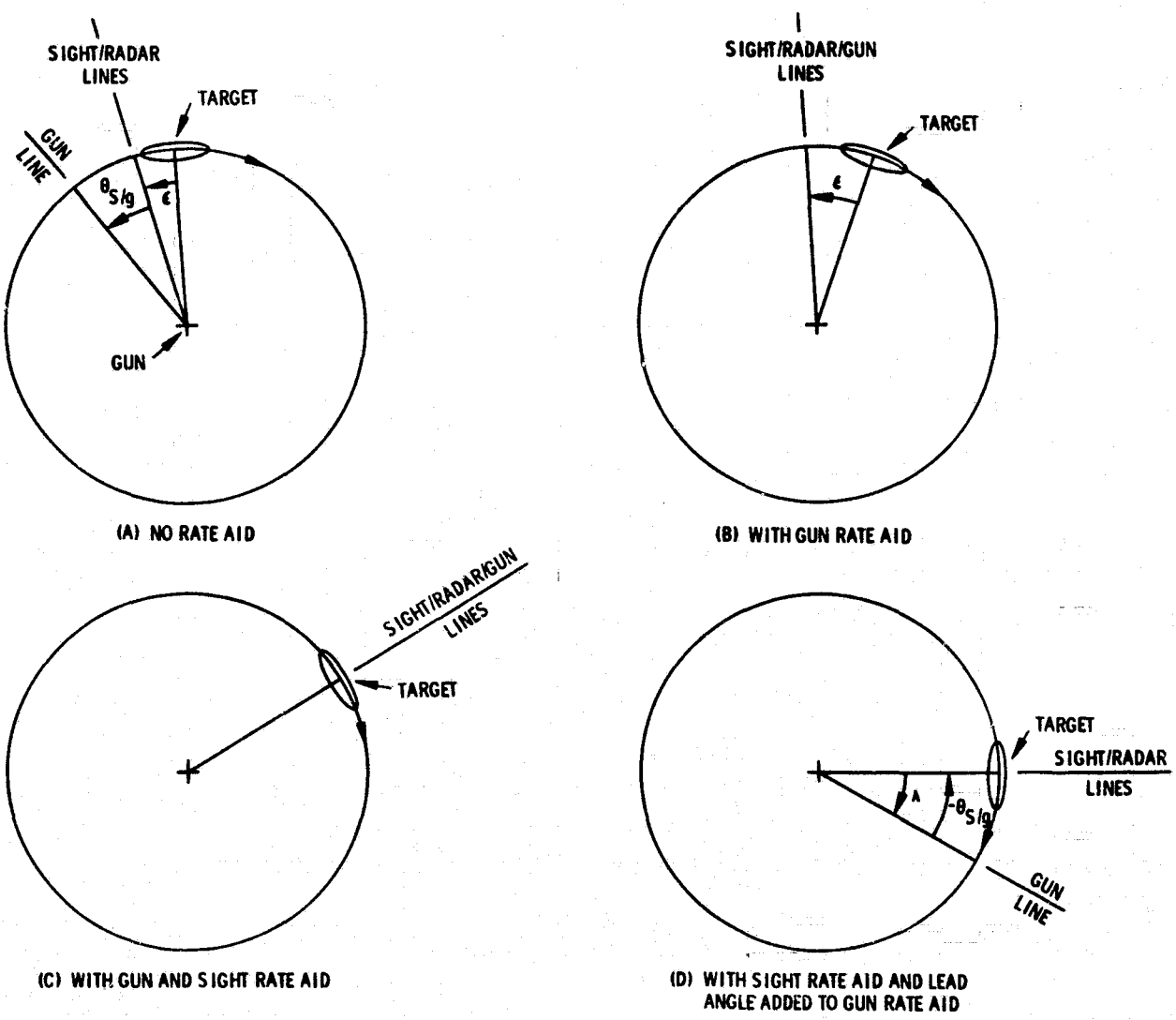


Figure 5. Tracking Operation Sequence (Azimuth) - Constant Velocity Target

Note, however, that the sight line (θ_S) and radar line (θ_A) are ahead of the gun line (θ_g). To achieve the zero lead angle, it is desirable to have gun line θ_g coincident with lines θ_S and θ_A. This is achieved by the computer. By operating on the system state data (figure 4) in the tracking computation, the computer outputs the +5 Vdc required at E_g by raising the value of V_g in figure 4. Now the gun turret starts to accelerate under the sight, but the gyro in the sight senses this and holds the sight line θ_S back on the target automatically. $\theta_{S/g}$ reduces during this process to zero causing the synchro transmitter (CX) output voltage to become zero. At this point, the computer is providing the +5 Vdc at E_g to move the turret at 0.5 radian/second; the operator/stick (track error sensor) is still providing the +5 Vdc at E_S in response to observed ε (note that rate gyro is moving in space and the output is -5 Vdc in response to turret velocity); and the relationship among the three lines is as shown in figure 5(b).

Now the computer outputs the +5 Vdc rate aid required at E_g by the rate gyro in order to eliminate the need for a tracking error ε. So as V_S is raised to 5 Vdc, the reticle starts to move closer to the target, ε goes to zero, and the error sensor stick output also goes to zero. With both rate aids applied, all three lines are coincident with the target line as shown in figure 5(c). The computer is now tracking and the human error sensor is merely observing the quality of the computer's track. If small corrections are required, the human supervisor can add corrective signals via the stick at the point E_e in figure 4.

Note that at this stage of operation, shown in figure 5(c), the operator is decoupled from the track loop and the computer is performing the continuous tracking function. The operator "communicates" with the computer via point E_e in figure 4. If the target is maneuvering, the operator adjusts the stick signal to correct track and the computer alters its target model to suit.

Still in control, the computer now determines the required lead angle λ via its ballistics computation. For example, it raises V_g from 5 Vdc to 7 Vdc in order to achieve "2 volts of lead angle" between the sight line θ_S and the gun line θ_g. Once again the gun turret accelerates under the sight and ahead of it because the rate gyro loop holds the sight line θ_S on the target. This process continues until the situation shown in figure 5(d) reaches steady state. At this point the gun line θ_g leads the sight line θ_S by lead angle λ. This means $\theta_{S/g} = -\lambda$ and the sight produces a negative CX signal from the sight of -2 Vdc which is summed with V_g = +7 Vdc to produce the steady state +5 Vdc required to move the turret at 0.5 radian/second.

In contrast to the complex visual/motor tasks that must be performed for the disturbed-line-of-sight (DLOS) system, the operator does nothing to produce the lead angle in the director system. The human role is that of mentor, monitoring how well the computer performs its task and providing corrective refinements to the track operation as required.

In presenting this explanation, a sequence of discrete steps was described. In actual practice, the operation is continuous and all functions change simultaneously. The concept of the operator as a tracking

error detector should be mentioned. The tracking error sensor could be: (1) an RF sensor if a radar receiver detector is used, (2) an electro-optical sensor if an IR or TV video tracker is used, or (3) a biological receiver detector if an operator is used. In figure 4, the error sensor could be any one of these. The Sharpshooter now in production uses the radar in one tracking mode and the operator in the optical tracking mode.

MICROPROCESSOR IMPROVES MAINTAINABILITY

All control system functions shown in figure 4 are performed within a microprocessor controller. The microcontroller communicates directly with the microcomputer and forms a powerful distributed processing system. Taking advantage of this capability provides system diagnostics, enhances maintainability, and reduces human-induced errors.

"...Human error accounts for at least 50% (underline added) of the failures of major [military] systems."* "The increasingly complicated nature of modern military systems together with shortages of qualified military personnel suggest that human-induced errors both in operation and maintenance of systems will increase unless more attention is given to this problem in the design and development phases of the acquisition process." "The problem of human-induced failures may very well become worse. Attendant to the increasingly complicated nature of systems are the lower education and aptitude levels of personnel now entering the services, the shortages and high turnover rate of experienced personnel, which lead to very low overall experience levels, and the effect of greater use of complex/sophisticated automatic checkout and built-in test equipment [that is difficult for the maintainer to use]."

One way of maintaining increasingly complex equipment while the level of skills of technicians decreases is to provide equipment for fault location and diagnosis that does not require highly skilled operators. The microprocessor that allows for the increased accuracy of the system described above also provides built-in test and diagnostic functions for system maintenance that would not be possible otherwise.

Both static and dynamic tests are conducted and, in many instances, faults are isolated to the circuit board level. Fault indications permit the operator to deal directly with some problems and refer maintenance personnel to areas the operator is not equipped to handle.

The built-in test equipment (BITE) philosophy automates BITE functions to the greatest degree possible minimizing operator participation in the test function. The computational capabilities of the microprocessor are used as much as possible to test the system rapidly. Tests are made to as basic a level as possible, and failure indications are correlated with the required maintenance test support equipment. This makes it unnecessary for test personnel to have to decide which tests to run. Computer automated static and dynamic tests are performed in addition to operational tests of the system in normal operating modes.

*This quotation and the others in this paragraph are all taken from Reference 8, p. 27.

During static tests power supply voltages and radar signals are sampled and compared with acceptable levels stored in memory. Coded error data are displayed on the control/display panel.

A basic operability test of the antenna and gun/turret servos is conducted in a similar way. More detailed diagnostic tests of drive signals to the optical sight, radar antenna, and weapon servo loops are also made in various combinations to ensure the proper functioning of those subsystems.

Through these aids the condition of the system can be assessed rapidly by relatively unskilled personnel. When maintenance action is required, the diagnostic capability provides enough guidance for a technician who is not highly skilled to be able to repair the system quickly.

CONCLUDING REMARKS

The improvements made in the manned-director type fire control system show large gains in human factors aspects as well as in system accuracy. Complex visual/motor tasks usually performed by the operator are now reduced to functions executed by a microprocessor, thus freeing the operator to oversee system operation. The microprocessor can also be programmed to test the system at regular intervals, which allows early location and diagnosis of faults. The built-in test feature enables maintenance personnel to circumvent repetitious checkout and troubleshooting procedures and repair the equipment using programmed instructions commensurate with their skills.

Future efforts in this area will be directed toward adaptive operator control filters, refinements in the built-in test programs, and toward adopting the console to increase the operator's effectiveness as supervisor of system performance.

REFERENCES

1. Livingston, W.L.: "Hold the Control Loop Design: Do the Systems Engineering First." Control Engineering, Jan 1981, pp. 79-82.
2. Sheridan, T.B.; and Ferrell, W.R.: Man-Machine Systems: Information, Control, and Decision Models of Human Performance. MIT Press, 1974.
3. Poulton, E.C.: Tracking Skill and Manual Control. Academic Press, 1974.
4. McCormick, E.J.: Human Factors in Engineering and Design. McGraw-Hill, 1976.
5. Kelley, C.R.: Manual and Automatic Control. Wiley, 1968.
6. Fogel, L.J.: Biotechnology: Concepts and Applications. Prentice-Hall, 1963.
7. Rawicz, H.C.; Bigley, W.J.; Cangiani, G.L.; and Yohannon, R.C.: Automated Fire Control Apparatus. U.S. Patent No. 4,004,729, Issued Jan. 25, 1977.
8. Effectiveness of U.S. Forces Can be Increased Through Improved Weapon System Design. U.S. General Accounting Office Report No. PSAD-81-17, Jan. 29, 1981.

D51

A THEORY OF HUMAN ERROR*

Duane T. McRuer, Warren F. Clement, and R. Wade Allen

Systems Technology, Inc.

SUMMARY

Human errors tend to be treated in terms of clinical and anecdotal descriptions, from which remedial measures are difficult to derive. Correction of the sources of human error requires that one attempt to reconstruct underlying and contributing causes of error from the circumstantial causes cited in official investigative reports. A comprehensive analytical theory of the cause-effect relationships governing propagation of human error is indispensable to a reconstruction of the underlying and contributing causes. This paper highlights a validated analytical theory of the input-output behavior of human operators involving manual control, communication, supervisory, and monitoring tasks which are relevant to aviation, maritime, automotive, and process control operations. This theory of behavior, both appropriate and inappropriate, provides an insightful basis for investigating, classifying, and quantifying the needed cause-effect relationships governing propagation of human error.

INTRODUCTION

Human error is of major concern in the development and deployment of man/machine systems. Human error is a significant contributing factor in aviation, maritime, automotive, and process control accidents. Thus the alleviation in number and consequence of human errors should be a primary goal of man/machine systems research. Traditionally, however, human error has been treated only tangentially. The measurement of task or system errors has routinely been employed in man/machine studies as a performance metric in the evaluation of other variables (e.g., equipment design, training, etc.). Human error has also been used in clinical and anecdotal terms as a convenient classification in accident investigations. Developing remedial measures from these applications is difficult, however, as errors have not always been classified according to a consistent structure; and other contributing factors or prevailing conditions have not been noted.

Recent research focusing directly on the nature and classification of human errors is changing the above state of affairs, however. Singleton (Refs. 1 and 2) has reviewed classification schemes, analytical techniques, and psychological theories in the study of human error. More recently Norman (Refs. 3 and 4) has been investigating applied human information processing and has evolved an action theory which he has used in the

---

* This research was sponsored by the Man-Vehicle Systems Research Division, Life Sciences Directorate, Ames Research Center, National Aeronautics and Space Administration under Contract NAS2-10400.

classification of errors made by highly-skilled operators in complex, high-demand systems. Most recently we have finished a report for NASA (Ref. 5) in which several behavioral models were reviewed for use in subsequent studies of human errors in aviation operations. These models cover continuous and discrete control, supervisory control, monitoring, and decision making and provide a basis for diagnostic investigation as well as research.

Human error is a complex, multifaceted phenomenon. In accounting for human error in complex man-machine systems we must consider both the spontaneous errors or "slips" delt with in Norman's action theory (Ref. 4) and more rational errors (having an assignable cause in hindsight) which arise due to problems in detection, perception, recognition, and judgment. The distinction here is that in one case the spontaneous error is seemingly aberrant and unintentional, whereas other errors can presumably be rationalized with behavioral theories that account for perception, judgment, decision making, monitoring, detection and recognition, and manual control.

DEFINITIONS

In previous work by Beek, et al. (Ref. 6) human error has been defined as an inconsistency with a *pre-defined behavioral pattern* established by virtue of system requirements and specifications and the design of the equipment and procedures to meet those specifications. This is a practical operational definition; however, it should be noted that incidents and accidents can arise because of inadequacies in the *design* of equipment and procedures. Errors may also be precipitated by environmental stress (physiological and psychological) impinging on the human operator. This has led us to differentiate between the *sources* and *causes* of human error. *Sources* are internal to the human operator and their consequences should be measurable as changes from normal or ideal human behavior which is consistent with system requirements. *Causes* are external factors which induce undesirable deviations in human behavior, such as unexpectedly large or extreme disturbances, high workload, distractions, inaccurate or noisy information, illusions, equipment design deficiencies, and inadequate training.*

Accompanying the current trend towards increasing automation in man-machine systems, there is increasing concern for errors induced by man-machine interaction (Ref. 7). In some cases errors are induced by increased complexity — the man-machine interface — and in other cases the operator's less active role as a monitor and supervisor seems to be the problem because there is a degradation of skill. At issue here is what the optimal level of operator involvement should be and the structuring of this involvement in order to minimize the occurrence and influence of human error on system performance.

---

* There could, of course, be internal causes of human error such as psychophysiological or neurological impairments. These should be handled with proper selection and periodic screening procedures which are not of direct interest here.

Errors or mismatches between desired and actual system or subsystem outputs are essential in situations where feedback is involved as an operating principle. Most of the time human operators use these errors to advantage in performing as error-correcting rather than error-avoiding system elements. For this reason in operations involving aviation, maritime, and automotive traffic control and process control, the errors per se are of major concern only when they are undesirable because of their size, timing, or character. These errors, which are intolerable in one way or another, we shall call *grievous errors*. In general, a *grievous error will involve an exceedence of safe operating tolerances*.

Human errors that do not always result in grievous errors may be nearly impossible to measure in practice unless behavioral identification techniques are employed. Behavioral identification may be performed by qualified observers (Refs. 8, 9, and 10) or by signal correlation analysis which can partition human error into coherent and incoherent components. Such identification of human errors which may be inconspicuous in one situation is very important, for they may lead to grievous errors in other circumstances. Thorough analyses of mission phase behavior sequences, both normal and abnormal, are necessary prerequisites to the application of behavioral identification techniques in the study of human error. Before considering some of the sources and causes of human error, we shall discuss the buildup of mission phase behavior sequences from constituent task behavior.

BUILDUP OF MISSION PHASE BEHAVIOR SEQUENCE(S) FROM CONSTITUENT TASK BEHAVIOR

A Perceptually Centered Viewpoint for Task Behavior

For a particular task the human component(s) as input-output elements consist of one (or more) of the pathways illustrated by Fig. 1 for one among several human operators of a system. Here the system inputs and errors may appear in several sensory modalities, and the motor subsystem output may be manipulative or verbal. The pathway used in a particular circumstance is the result of the nature of the perceptual field and of training. Table 1 summarizes these and other facets of this perceptually centered model of human behavior.

The human's operations are thus defined as an open-loop, closed-loop, or open- and closed-loop behavior pattern with identified sensory input and motor output modalities. For some inputs, of course, there is no immediate output; instead, the information received may simply be stored in memory. In other cases the lack of a measurable output should nonetheless be interpreted as the 0 portion of a 0,1 binary pair of possibilities.

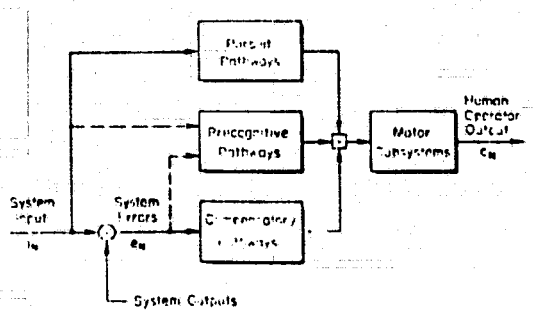


Figure 1. Three Paths in Perceptually Centered Model of Human Behavior

TABLE 1. SUMMARY CHARACTERISTICS OF PATHWAYS IN PERCEPTUALLY-CENTERED MODEL OF HUMAN BEHAVIOR

PATHWAYS SELECTED	PERCEPTUAL FIELD CONTENT	ACTION OR OUTPUT	CORRELATES OF TRANSITION AMONG LEVELS OF SOP
Compensatory	Narrow; deviations only	Designed to correct exceedences and reversals; not necessarily rehearsed	Expanding perceptual field content and extension thereof; Increasing time delay; Increasing bandwidth; Increasing rehearsal; Increasing perceptual motor load
Pursuit	Broader; separable inputs, outputs, commands, disturbances in addition to deviations	Designed to correct deviations and to compensate for internal delay; moderately well rehearsed	
Precognitive	Exceedingly broad and extended, even among other individuals and organizations by means of a conference, by recall of past experience, or by recruitment of other resources; separable inputs, outputs, commands, disturbances only; feedbacks not necessary	Discrete; cued, transient; very well rehearsed	

Monitoring and Decision Making Viewpoint for Task Behavior

With increased use of automatic controls and computers in modern day aircraft, traffic, and process control systems, the role of the human operator is becoming more supervisory, involving increased amounts of monitoring and decision making. In these roles, human outputs are typically discrete (as opposed to continuous control actions) and include verbal communication as well. Monitoring and decision making errors can arise due to misperception of monitored information and misinterpretation of perceived information. Errors can also occur in the more cognitive aspects of decision making where the operator must account for various possible consequences of the alternative actions available to him.

Monitoring and decision making constructs and viewpoints are useful in several ways. First, human errors sometimes appear to be inexplicable when, for example, only two courses of action are possible, and an operator appears to make the obviously wrong choice. By considering the elements of these task situations in a decision making context, one can gain additional insight into the underlying factors involved. Second, if specific analytic decision-making models are appropriate descriptors of the mission phases being examined, then the model can serve as a means for the analysis and interpretation of the operational or experimental results. Third, a combination of monitoring, decision making, and control viewpoints is essential in treating repeated trials in an experiment or an ensemble of simulations involving many crews. In a single trial, behavior and performance for all the tasks involved are specific concrete actions (or inactions) flowing in a sequence. Error is identified as an extreme deviation from a desired state. Among many trials these concrete actions often exhibit differences, either in kind or in degree. A probabilistic structure for particular events then becomes appropriate as a means of describing the experimental data. Further, the potential tradeoffs (based on experience and training) involved in selecting various emergency actions

can be exposed in the light of a utility concept. Monitoring and decision making theories are the appropriate vehicles for such considerations.

Partition of the Mission into Phases, Tasks, Skills, and Outcomes

If we are to apply these elementary behavioral models to complex operations of men and machines, they must be associated with sequences of operations which, together, serve to accomplish a desirable end, i.e., a mission. To accomplish this the mission is first defined and partitioned into a hierarchy of constituents. The primary constituents are *mission phases*. These are of a size and duration which allow the broadest factors (e.g., environmental variables) that influence human behavior to be identified. At the next level are *tasks*, which are associated with a particular operation in a sequence and are sized to permit the identification of "critical" skills. Aberrations in the execution of these skills ultimately determine the sources of contributions to human error.

A mission phase may be broken down into various subdivisions depending upon its complexity. For our purposes here we are ultimately interested in the elemental unit of all phases involving the human operator, the *task*. As a working definition here we will define a task as *an activity at the functional interface of the human operator and the individuals, objects, and environments with whom or which he interacts* (adapted from Ref. 11). We will further specify a task for our purposes here as a goal- or criterion-oriented work increment involving application of a *skill* or *set of skills* by the human operator. Thus, by partitioning the mission phases into tasks, we can then identify those fundamental human operator behavioral factors, *skills*, which influence operational safety. For tasks which are critical to safety (i.e., exert a predominant influence in some sense), it is the *proficiency* with which a skill or set of skills is applied that we wish to consider in order to identify the underlying sources of human error.

In preparing the operations breakdown for a particular mission phase, each task for each operator is listed as an item in an ordered, nominal sequence. Conceivably this order might be changed or omitted in off-nominal circumstances, and this by itself may be a cause of error. Associated with each task are input and output modalities for each operator in his respective relationships with other operators and equipment. Associated also with each task is an indication of the human behavior characteristics nominally involved in carrying out the task at hand. In many cases the nominal behavioral characteristics may not be exhibited by actual operators, and abnormal behavior may result in an out-of-tolerance system error.

For the study of human error, the nominal task breakdown must therefore be further subdivided to account for all possible *outcomes* induced by abnormal behavior. In this endeavor the application of Murphy's law and its corollaries can be helpful. Other off-nominal aspects which should be considered are the accumulation of stress and degradation of skill. Each mission phase presents a combination of *environmental* and *task* stresses on the operators, and these stresses influence operator performance. After lapses in operational practice or long intervals of inactivity, individuals have to cope with the problem of maintaining proficiency of skills which may be critical to safety. Skills performed *infrequently*, for whatever reason, are most likely to fall into this category. Of these skills, those having

high workload factors by virtue of being time-constrained or because they involve complex operations are most likely to cause serious performance decrements. Several conditions may contribute to the degradation of these skills: (a) lack of practice, (b) inability to practice in the appropriate environment, (c) interference or negative transfer arising from the practice of competing skills, and (d) physiological deconditioning due to fatigue induced by the environment or due to alcohol or drug stresses. The tasks which are most likely to be affected by these human conditions should be especially flagged for investigation.

In most of the tasks where precognitive operations are identified as nominal or customary, additional qualification is necessary. Such open-loop operations are normally of limited duration and are properly interspersed or concluded with closed-loop operations either directly, as in dual mode continuous control, or indirectly in the context of an off-line supervisory monitor. Omission of the closed-loop monitoring activity may in fact lead to human error as shown in Ref. 12. To examine the role of a supervisory monitor in more detail, we next consider some models for the integration of the three functional pathways in Fig. 1.

INTEGRATION OF THE PATHWAYS — THE METACONTROLLER

Each pathway in Fig. 1 contains a number of subsets of behavior appropriate to the task. Assume that identifiable prerequisite conditions and limits can be found (e.g., experimentally) for each subset of observed behavior. Then one model for the perceptual organization process would be an active off-line supervisory monitor which identifies the conditions that currently exist, selects and activates some most likely pathway/subset, monitors the result, reselects a new pathway/subset when necessary or when further information is identified as a result of the first operations, and so forth. Appropriately this has been termed the metacontrol* system by Sheridan in Ref. 13. A simplified diagram of such a metacontroller is given in Fig. 2a. Other preliminary work on an algorithmic-type model for the successive organization of perception (SOP) process is given in Ref. 14. The possibilities for error due to inappropriate activities within such a system are manifold. Such a model provides a logical basis for understanding some of the causes underlying selection of an inappropriate behavioral model which may ultimately lead to an identifiable error.

An appropriate form for this model is a flow or decision process algorithm. Related models have been described in Refs. 16 and 17, and applied to a specified task involving a given sequence of subtasks in Refs. 18 through 21. Thus the algorithmic approach is by no means novel. Most of these attempts have had limited application because of the inordinate complexity and repetitive cycling required to represent continuous tasks. Yet by breaking out the compensatory and pursuit pathways as separate entities which handle most of the continuous operations, the metacontroller of

---

* Metacontrol = the human's activity-supervising control, transcending the various directly involved systems such as the perceptual, central, and neuromuscular systems (from Greek "meta" meaning "involved with changes").

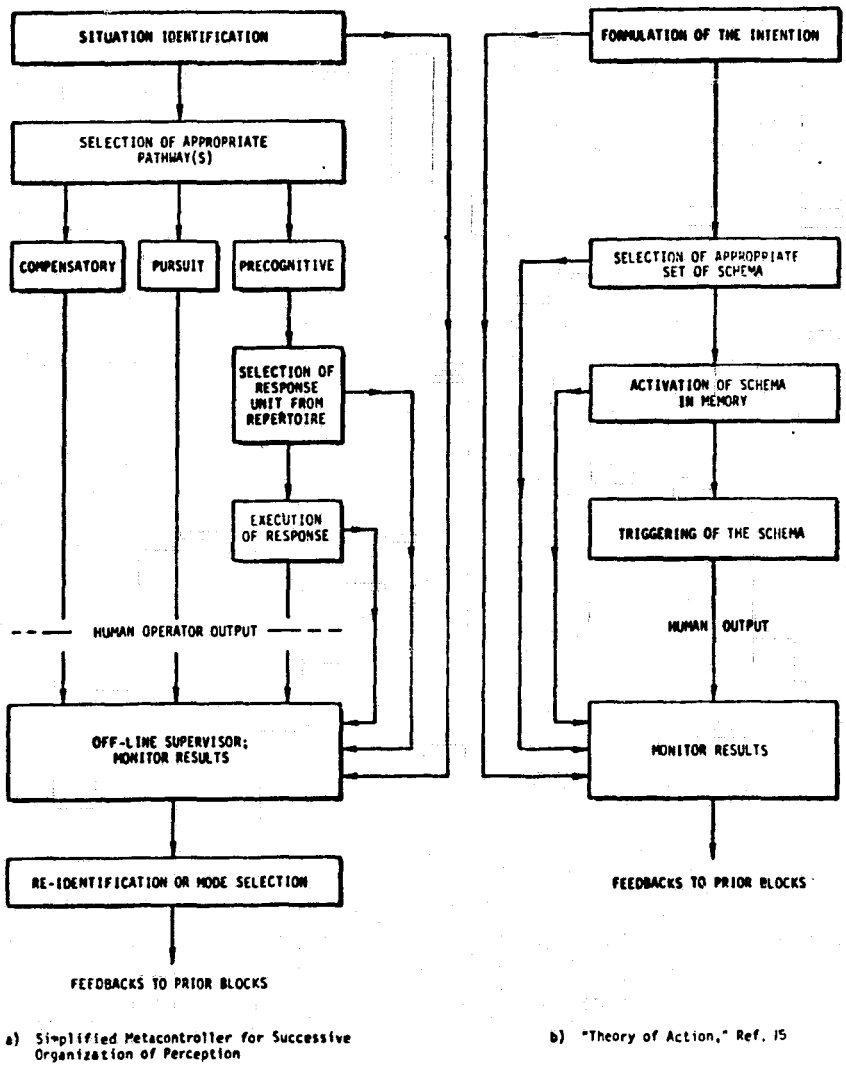


Figure 2. Flow Diagram for SOP Operations and the Ref. 15 Theory of Action

Fig. 2a gets around some of these problems. Algorithmic models are used where they are best suited (logical functions), while isomorphic models of human behavior are used where they are most efficient (well-defined tracking or stimulus-response situations). Continuing research in the disciplines of observation, pattern recognition, estimation, and timeshared processing should yield additional material useful to the interpretation of SOP. For example, Noton offers a sequential pattern perception and recognition theory in Ref. 22 which appears to have connections with SOP and other models which have been found useful in characterizing human behavior.

A particularly interesting parallel to the SOP metacontroller which is especially valuable for the understanding of error is given in Ref. 15. The "Theory of Action" proposed there has a number of cognitive stages and components. The base stores for action are organized memory units or sensorimotor knowledge structures —— "schemas" which control skilled action sequences. A basic control sequence starts with intention, and proceeds through selection, activation, and triggering of schema to result in an output action. The results at various levels in this sequence are monitored, and may be modified by feedbacks to the previous stages. A simplified block diagram for this theory is shown in Fig. 2b. Its elements are clearly similar to the precognitive elements in the metacontroller of

Fig. 2a. Much of the Fig. 2b model is based on the study of verbal "slips," which can be errors by another name; so the connections between human manual control and verbal activities are very useful in our search for generalization.

Using this overall structure as a point of departure, we progress in the next topic to discuss some sources and causes of human error.

SOURCES AND CAUSES OF HUMAN ERROR

The functional pathway triad and metacontroller model for human behavior illustrated in Fig. 2a contains within its structure many features which can, in abnormal versions, lead to grievous system errors. These features we shall refer to as *sources* or *antecedents* of error. Sources are endogenous or internal to the human. Their consequences are all measurable in terms of changes from ideal or nominal human behavior for a particular task. These changes may be induced by external (exogenous) factors which will be referred to as *causes* of errors. The first two columns of Table 2 illustrate these distinctions for compensatory operations.

The remaining two columns of Table 2 present a verbal synthesis of a great deal of empirical data from many experimenters. All of the currently demonstrated forms of abnormal compensatory input-output behavior are represented here. In total they represent an error source which can be described generally as *inappropriate perception, decision, and/or execution within a selected level (in this case, compensatory) of organization of behavior.* The sources of error in this framework are summarized in Table 3.

In principle tables similar to Table 2 can be constructed for the other source possibilities in Table 3, e.g., Table 4 for pursuit operations. However the experimental data base for most of these is nowhere near as comprehensive as it is for the compensatory pathway. Many of the elements in the precognitive pathway can be developed, by analogy, from Table 1 of Ref. 15, which lists the presumed sources of "slips" (or errors) in the structure of Fig. 2b.

Transitions from higher to lower levels of skill occur when the attentional field becomes too narrow. They can also occur when the human is sufficiently impaired perceptually (i.e., by alcohol, fatigue, hypoxia, etc.) so that action as a multi-channel operator is significantly degraded. In these instances divided attention is possible only by switching to and fro as an essentially single channel information processing device.

Although probably one of the most fundamental sources of human error, the inappropriate *organization* of perception and behavior for the task at the executive level of the metacontroller has received much less attention in the literature than have inappropriate perception, decision, and/or execution *within* a selected level of behavioral organization. The SOP theory described in Ref. 23 offers a unifying approach to inappropriate *organization* as a source of human error.

TABLE 2. BEHAVIORAL SOURCES OF ERROR IN COMPENSATORY SYSTEMS

SINGLE CHANNEL OPERATIONS

BASIC SOURCE (ENDOGENOUS)	CAUSES (EXOGENOUS)	OPERATOR BEHAVIOR	EFFECTS ON SYSTEM
Extreme command or disturbance amplitudes	Unexpectedly large command or extreme environment	Operator response normal	System overloaded, forced out of tolerance although operating properly
Extreme command or disturbance bandwidth	Broadband input signal noise; Unexpectedly broadband disturbance	Regression of crossover frequency	Reduced system bandwidth
Controlled-element change	Mulfunction/failure in controlled element	Affecting output for transient interval; Adaptation to new controlled element	Transient errors during transition; Reduced system bandwilth
Reduced attention field	Poor signal/noise ratio (e.g., poor contrast, high intensity distraction stimuli, low level signals, etc.)	Operator threshold, net gain reduction	System bandwidth reduction; (missed signals as one extreme)
Reversals	Misperception of error sign; Naivety	Remnant increase; Intermittently reversed output	Increased system noise; Intermittently reversed system output

MULTI-INPUT OPERATIONS

BASIC SOURCE (ENDOGENOUS)	CAUSES (EXOGENOUS)	OPERATOR BEHAVIOR	EFFECTS ON SYSTEM
Divided attention, perceptual scanning	Increased informational requirements for monitoring or control	Remnant increase (scanning); Increase in loop gains; Simultaneous multi-channel operations	Increased system noise; Reduced bandwidth
	Information overload: Too many separate input channels; Too many significant signals; Backlog of unattended operations	As above, plus failure to detect some signals, increased latencies, and missed output responses	Saturation; Missed responses; Instability in the mean square sense
Reduced attentional field	Operator impairment (fatigue, alcohol, hypoxia, etc.)	Remnant increase over scanning; Further decrease in loop gain; Sequentially-switched single channel operations; Deletion/missed responses	Increased system noise Reduced bandwidths Increased latencies Missed responses
Illusions, kinetosis	Conflict between or among visual, vestibular, aural, kinesthetic and/or proprioceptive inputs	Remnant increase; Decrease in operator's gain; Mal a propos responses; Missed responses	Increased system noise Reduced bandwidth Mal a propos responses Missed responses

CONCLUSIONS

The input-output behavior of human operators in manual control systems is characterized by an internal organization involving three major pathways. These correspond to closed-loop, combined open- and closed-loop, and open-loop behavior patterns. In manual control systems which exemplify these patterns, the system bandwidths, attentional fields, and rehearsal requirements are ordered correspondingly, i.e., compensatory < pursuit < precognitive. Similar but inverted orderings of perceptual motor loading and system latencies are associated with the three pathways.

TABLE 3. SOURCES OF HUMAN ERROR

(Sources are endogenous or internal to the human operator by definition)

Inappropriate perception, decision, and/or execution within a selected level of behavioral organization

- **Compensatory (expanded in Table 2)**
- **Pursuit (expanded in Table 4)**
- **Precognitive (expanded in Table 1 of Ref. 15)**
  - Selection of response unit
  - Execution of response

Transitions from a higher to lower level of behavioral organization

- **Precognitive to pursuit**
- **Precognitive to compensatory**
- **Pursuit to compensatory**

Inappropriate organization of perception and behavior for the task at the executive level of the metacontroller

(Items 1-5 are associated with the "situation identification" block in Fig. 2a)

Errors in:

- (1) Formulation of intent, assignment of function and its priority
- (2) Identificatior of specific task/situation/action: continuous or discrete
- (3a) Selection of likely sources of information and their temporal order (i.e., past, current, or preview)
- (3b) Assignment of priority in sources of information among inputs and feedbacks
- (4) Identifying predictability or cohernece in and among sources of information
- (5) Identifying familiarity with the task

(Item 6 is associated with the "selection of appropriate pathway(s)" in Fig. 2a)

- (6) Organizing operation on inputs and feedbacks.

Inadequate off-line monitor/supervisor in the metacontroller

The three-pathway model for manual control can be generalized to a perceptually-centered model appropriate for input-output human behavior involving sensory modalities other than vision and output modalities other than manipulation.

The perceptually-centered model for human behavior is further generalized to include an executive and supervisory-monitoring metacontroller which identifies the situation, selects the appropriate pathway, directs the information flow through the pathway selected, and monitors, on an off-line basis, the resulting outputs. The off-line monitoring feature constitutes yet another feedback, albeit on an intermittent and longer term basis.

The characterization of human behavior presented here provides a rational basis for planning specific investigations of the sources of human error, either for the purpose of research in advance or diagnosis after the fact. When the purpose and scope of an investigation has been set forth, the behavioral models summarized here can be used to predict (sometimes), subsume, describe, and rationalize the experimental or operational results.

TABLE 4. BEHAVIORAL SOURCES OF ERROR IN PURSUIT OPERATIONS
(Multi-Input Operations, by Definition)

BASIC SOURCE (ENDOGENOUS)	CAUSES (EXOGENOUS)	OPERATOR BEHAVIOR	EFFECTS ON SYSTEM
Controlled element change	(see corresponding causes in Table 2)	Transient regression to compensatory level (see corresponding behavior in Table 2)	Transient errors during transition; Reduced system bandwidth
Divided attention, perceptual scanning	(see corresponding causes in Table 2)	Remnant increase; Decrease in operator's gain; (see also corresponding behavior in Table 2)	Increased system noise; Reduced bandwidth; (see also corresponding effects in Table 2)
Reduced attentional field in spatial dimensions	Poor input and/or error signal/noise ratio (e.g., inability to identify input. Task involves disturbance regulation rather than command-following and disturbance cannot be identified; Mismatched scaling between input and error; Distortion of input; Lack of input conformability with visual field; See also corresponding causes in Table 2	Remnant increase; Operator's threshold on input may cause missed responses and regression to compensatory level; Operator's threshold on error may reduce gain in or open compensatory loop (see also corresponding behavior in Table 2)	Increased system noise; Reduced system bandwidth (missed responses as one extreme)
Reduced attentional field in temporal dimension, i.e., reduced preview	Inability to identify <u>future</u> input or disturbance; Prodigious extrapolation required to estimate <u>future</u> input or disturbance	As above, plus increased latencies	As above, plus increased response latencies
Reversals	Perceptual inversion of input; Faulty input-background discrimination; Lack of input conformability with visual field	Remnant increase; Intermittently reversed output	Increased system noise; Intermittently reversed output
Illusions, kinetosis	(see corresponding causes in Table 2)	Remnant increase; Decrease in operator's gain; Mal a propos responses; Missed responses	Increased system noise; Reduced bandwidth; Mal a propos responses; Missed responses

REFERENCES

1. Singleton, W. T. "Theoretical Approaches to Human Error," *Ergonomics*, Vol. 16, No. 6, November 1973 pp. 727-737.
2. Singleton, W. T., "Techniques for Determining the Causes of Error," *Applied Ergonomics*, Vol. 3, No. 3, 1972, pp. 126-131.
3. Norman, D. A., "Human Error," *Psychology Today*, January 1980.
4. Norman, D. A., "Categorization of Action Slips," *Psychological Review*, January 1981.
5. McRuer, Duane T., Clement, Warren F., and Allen, R. Wade, *A Theory of Human Error*, Systems Technology, Inc., Technical Report No. 1156-1 (forthcoming NASA Contractor Report), May 1980.
6. Beek, C., et al., *Human Reliability Research*, Technical Report 430, Operations Research, Inc., Silver Spring, Md., 1967 (Available from DDC as AD 664495).

7. Wiener, E. L. , and Curry, R. E., *Flight-Deck Automation: Promises and Problems*, NASA TM 81206, June 1980.
8. Sanders, A. F., "Some Remarks on Mental Loads," in Neville Moray (Ed.), *Mental Workload. Its Theory and Measurement*, Plenum Press, New York, 1979, pp. 41-77.
9. Lee, Alec M., *Applied Queueing Theory*, Macmillan, London, 1966.
10. Ruffell-Smith, H. P., *A Simulator Study of the Interaction of Pilot Workload with Errors, Vigilance, and Decisions*, NASA TM 78482, 1979.
11. Miller, R. B., "Task Description and Analysis," Chapter 6 of *Psychological Principles in System Development*, 1962, pp. 187-228.
12. Sanders, A. F., "Some Aspects of the Selective Process in the Functional Visual Field," *Ergonomics*, Vol. 13, No. 1, 1970, pp. 101-117.
13. Sheridan, T. B., "The Human Operator in Control Instrumentation," in R. H. Macmillan, et al., (Eds.), *Progress in Control Engineering*, Academic Press, New York, 1962, pp. 141-187.
14. McRuer, D. T., Hofmann, L. G., Jex, H. R., Moore, G. P., Phatak, A. V., Weir, D. H., and Wolkovitch, J., *New Approaches to Human-Pilot/Vehicle Dynamics Analysis*, AFFDL-TR-67-150, 1968.
15. Norman, D. A., *Slips of the Mind and an Outline for a Theory of Action*, University of California, San Diego, CHIP 88, 1979.
16. Newell, A., Shaw, J. C., and Simon, H. A., "Elements of a Theory of Human Problem Solving," *Psychol. Sci.*, Vol. 4, 1959, pp. 161-166.
17. Reitman, W. F., "Heuristic Programs, Computer Simulation and Higher Mental Processes," *Behavioral Sci.*, Vol. 4, 1959, pp. 330-335.
18. Thomas, R. E., and Tou, J. T., "Human Decision-Making in Manual Control Systems," in *Second Annual NASA-University Conference on Manual Control*, NASA SP-128, 1966, pp. 325-334.
19. Siegel, A. I., and Wolf, J. J., "A Technique for Evaluating Man-Machine Systems Designs," *Human Factors*, Vol. 3, 1961, pp. 18-28.
20. Braunstein, M. L., Laughery, K. R., and Seigfried, J. B., *Computer Simulation of Driver Behavior During Car Following: A Methodological Study*, Cornell Aeron. Lab., YM-1797-H-1, 1963.
21. Thomas, R. E., *Developments of New Techniques for Human Controller Dynamics*, Aerospace Med. Res. Labs., MRL-TDR-62-65, 1962.
22. Noton, D., "A Theory of Visual Pattern Perception," *IEEE Trans*, Vol. SSC-6, No. 4, 1970, pp. 349-357.
23. McRuer, D. T., and Krendel, E. S., *Mathematical Models of Human Pilot Behavior*, AGARDograph 188, 1974.

A Normative Simulation Model of AAA Crew/System Performance and Decision-Making

Sheldon Baron, Greg L. Zacharias, Ramal Muralidharan and Marcia P. Kastner

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, MA 02238

A normative model has been developed to simulate the continuous control performance and discrete decision-making activities performed by a AAA crew during an engagement. The model consists of three basic sub-models: a system model of the threat aircraft and AAA system hardware; a commander model which simulates the primary situation assessment and decision-making activities of the commander; and a gunner model which simulates the primary monitoring and continuous control activity of the gunner.

The system model provides options for simulating one or more countermeasures (CM's) used by the threat aircraft: evasive maneuvers, electronic CM's, or optical CM's. The model also allows for three different operating modes (full auto with radar tracking and linear target prediction, semi-auto with gunner sight tracking and linear prediction, and full manual with gunner tracking and lead prediction) to provide the commander with flexibility in countering the threat CM's.

The commander model is a supervisory control model extension of the Optimal Control Model (OCM), and simulates the situation assessment and decision-making activities required for a successful AAA engagement. Basic information-processing of the cues available to the commander is accomplished by a non-linear extension of the OCM estimator, which operates on spherical-frame "visual" displays to generate Cartesian-frame estimates of the target trajectory state. A "situation-assessor" then mixes these estimates with discrete auditory/visual "alarms", to provide situational probabilities of CM use, target location in the firing envelope, etc. Based on this internally perceived situation, the "decision-maker" then maximizes the expected net gain (ENG) for switching (or not) to an alternate operating mode. The mode-selection decision, along with appropriate firing strategy decisions, are then communicated to the gunner for subsequent action.

The gunner model is similar to past OCM gunner models, and simulates monitoring behavior, continuous tracking performance, event detection, and communication. The model adds to the OCM-based monitor/estimator/controller a provision for switching among different monitoring strategies, internal models (of the AAA system dynamics), and performance objectives, based on the desired system operating mode. Also included is an event detection capability, and verbal communications channels with the commander,

to relay relevant event occurrences and to follow operational commands.

Preliminary simulation of the crew/system model has been completed, over a number of scenarios and engagement sequences. A rich variety of metrics are available from the model (e.g., continuous tracking performance, intercrew message time-lines, mode decision activity, etc.) and it is anticipated that future work will concentrate on developing appropriate experimental scenarios for verifying specific predictions of the model.

STABILITY ANALYSIS OF AUTOMOBILE DRIVER STEERING CONTROL*

R. Wade Allen

Systems Technology, Inc.
Hawthorne, California

SUMMARY

In steering an automobile the driver must basically control the direction of the car's trajectory (heading angle) and the lateral deviation of the car relative to a delineated pathway. This paper considers a previously published linear control model of driver steering behavior which is analyzed from a stability point of view. A simple approximate expression for a stability parameter, phase margin, is derived in terms of various driver and vehicle control parameters, and boundaries for stability are discussed.

A field test study is reviewed that includes the measurement of driver steering control parameters. Phase margins derived for a range of vehicle characteristics are found to be generally consistent with known adaptive properties of the human operator. The implications of these results are discussed in terms of driver adaptive behavior.

INTRODUCTION

Analysis of the closed-loop dynamic behavior of the driver/vehicle system can give some insight into driver and model behavior, and provide simplified analytical expressions for relationships between various model parameters. Here we will use a linear, two degree of freedom vehicle model. The two degrees of freedom to be considered are heading, or direction of vehicle motion, and lateral position. These two variables are under direct control of the driver. A third basic vehicle mode not considered here is roll angle. Although roll angle may influence driver behavior, it is not controlled per se by the driver. The discussion here will include a summary of previously published work (References 1 and 2).

The two degree of freedom model should be considered as an equivalent or approximation to higher degree models. Thus it subsumes such things as roll steer and weight transfer effects to a first approximation, so that the heading response of the model is an adequate approximation to a real vehicle for similar inputs.

DRIVER/VEHICLE SYSTEM MODEL

A block diagram of the driver/vehicle system model is shown in Figure 1. The vehicle equations generate side velocity (v) and yaw (heading) rate as a function of steering inputs through the G_δ^v and G_δ^r transfer functions, respectively. Kinematic equations then compute vehicle heading angle and lateral lane position from side velocity and yaw rate inputs. The driver finally develops steering corrections based on perceived heading and lane position errors as processed by the behavioral transfer functions Y_y and Y_ψ.† For the closed-loop analysis reviewed here we will consider steering against disturbances applied at the steering point as shown in Figure 1.

---

*This work was partially funded by the Automotive Safety Affairs Office of the Ford Motor Company. However, the contents of this paper represent the views of the author and do not necessarily reflect the official views or policy of the Ford Motor Company.

†At this point a question might be raised as to why the Y_ψ block is not placed in the ψ_e pathway but rather is in the ψ_L pathway, which is the sum of the heading error and some function Y_y of lane position error y_e. This arrangement is consistent with the perceptual information most readily available to the driver, which is further described in Ref. 3.

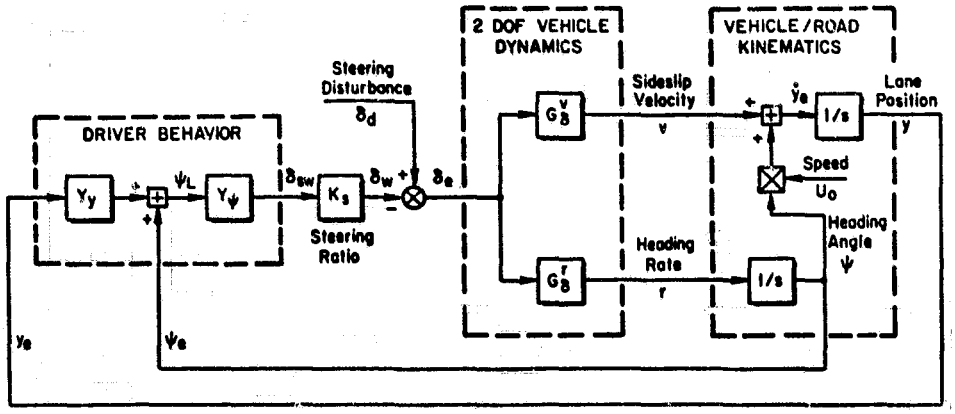


Figure 1. Driver/Vehicle Model with Two Degree of Freedom Vehicle Dynamics

This disturbance input will be used to approximate the lateral response effects of roadway disturbances on the wheels, and various forces and moments caused by roadway slope and lateral wind gusts.

The dynamics and stability of the Figure 1 system can best be analyzed by considering a steering disturbance signal (δ_d) as the input to a closed-loop system, and then analyzing the total open-loop transfer function between δ_e passing through the vehicle dynamics and driver behavior to the δ_w point. In this way we subsume the multi-path portions of the system as we progress from the single control input variable δ_e to the single control output variable δ_w.

Through simple block diagram algebra we can now derive an expression for the open-loop transfer function δ_e/δ_w as follows. First the heading angle and lane position errors cn be expressed in response to δ_e inputs as

$$\psi_e = G_\delta^\psi \delta_e \quad ; \quad y_e = \frac{\delta_e}{s}\left(G_\delta^v + U_o G_\delta^\psi\right) \tag{1}$$

Next δ_w steering response can be expressed in terms of perceived heading angle and lane position errors as

$$\delta_w = K_s \delta_{sw} = Y_\psi(\psi_e + Y_y y_e) \tag{2}$$

Combiing Equations 1 and 2 we can then express the total open-loop transfer function as

$$\frac{\delta_w}{\delta_e} = Y_\psi \frac{K_s G_\delta^r}{s}\left\{Y_y\left(\frac{G_\delta^v}{G_\delta^r} + \frac{U_o}{s}\right) + 1\right\} \tag{3}$$

This expression can be further simplified if we now express the transfer function between heading and lane position to control input in terms of the vehicle dynamics and kinematic equations:

$$G_\delta^\psi = \frac{\psi_e}{\delta} = \frac{G_\delta^r}{s}$$

$$G_\delta^y = \frac{y_e}{\delta} = \frac{G_\delta^v + (U_o/s)G_\delta^r}{s} \tag{4}$$

Combining Equations 3 and 4 we have

$$\frac{\delta_w}{\delta_e} = Y_\psi K_s G_\delta^\psi \left\{ Y_y \frac{G_\delta^y}{G_\delta^\psi} + 1 \right\} \tag{5}$$

Now consider models for each of the component transfer functions in Equation 5.

Vehicle Dynamics

Two degree of freedom vehicle dynamics have previously been analyzed in some detail (Reference 1) from which the following material has been summarized. In general the transfer functions between heading angle and lane position to steering control input can be expressed by second-order equations.

$$G_\delta^\psi = \frac{N_\delta (s + 1/T_r)}{s(s^2 + 2\zeta_1\omega_1 s + \omega_1^2)} \tag{6}$$

$$G_\delta^y = \frac{Y_\delta (s^2 + 2\zeta_y\omega_y s + \omega_y^2)}{s^2(s^2 + 2\zeta_1\omega_1 s + \omega_1^2)} \tag{7}$$

where T_r is the basic time constant of the vehicle's heading response and the remaining coefficients are functions of various vehicle parameters.

Using the few basic vehicle parameters described in Figure 2 and some simple assumptions we can now express Equations 6 and 7 in terms of vehicle characteristics as follows. The inverse of the heading time constant is given by

$$T_r^{-1} = \frac{2(a + b)}{mU_o a} Y_{\alpha 2} \tag{8}$$

If we now assume that the vehicle radius of gyration (k_z) is approximately equal to the geometric mean of the axle to c.g. distance:

$$k_z = \frac{I_{zz}}{m} = \sqrt{ab} \tag{9}$$

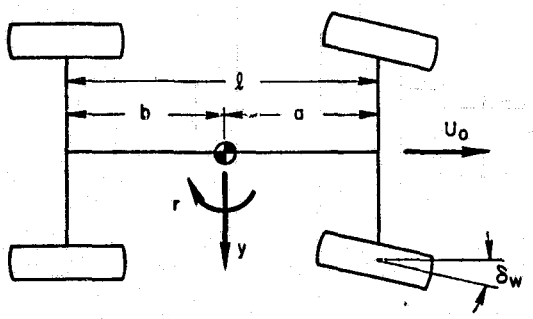


Figure 2. Vehicle Parameters for Steering Dynamics

and introduce two additional parameters, the stability factor K (sec^2/ ft^2), which is related to the SAE understeer/oversteer gradient (Reference 1)

$$K\ (deg/sec) = 1847(a + b)K \tag{10}$$

and the "axle load ratio"

$$\nu = \frac{bY_{\alpha 2}}{aY_{\alpha 1}} \tag{11}$$

we can then write approximate expressions for the remaining coefficients in Equations 6 and 7.

$$\begin{aligned}
&Y_\delta = \frac{2}{m} Y_{\alpha 1} \quad ; \quad N_\delta = \frac{2a}{mk_z^2} Y_{\alpha 1} \quad ; \quad \frac{Y_\delta}{N_\delta} = \frac{k_z^2}{a} \cong b \\
&\omega_1 \cong T_r^{-1} \sqrt{\frac{1 + KU_o^2}{\nu}} \quad ; \quad \zeta_1 \cong \frac{\nu + 1}{2\sqrt{\nu(1 + KU_o^2)}} \\
&2\zeta_1 \omega_1 \cong T_r^{-1} \left(\frac{1}{\nu} + 1\right) \\
&\omega_y^2 \cong T_r^{-1} \frac{U_o}{b} \quad ; \quad 2\zeta_y \omega_y \cong T_r^{-1}
\end{aligned} \tag{12}$$

At best, the above expressions are good approximations for many cars. At worst the expressions should give us a qualitative feel for the lateral dynamic characteristics that are of importance to the driver.

Inspection of the above equations can give us insight into vehicle dynamic response characteristics that are important from a driver control point of view. It is obvious that the heading time constant (T_r) dominates the vehicle dynamics, and that it is a direct function of speed (Equation 8). In Equations 6 and 7 the numerator and denominator roots are an inverse function of speed (i.e., T_r^{-1}), so as the vehicle increases in speed the heading response becomes slower (lower frequency). The stability factor K can also exert further influence as a function of speed. For an oversteering car ($K < 0$) the speed sensitivity of the heading mode is even further exaggerated, while the damping decreases, causing the car to become oscillatory. For an understeering car ($K > 0$) the speed sensitivity of the heading mode denominator is reduced and the damping increases with speed.

Another factor to consider is the heading rate sensitivity to steering inputs. If we evaluate the derivative of Equation 6 at zero frequency we end up with simple expressions for steady-state heading rate and side acceleration:

$$\begin{aligned}
G_\delta^r{}_{s=0} &= U_o/\ell(1 + KU_o^2) \\
G_\delta^{a_y}{}_{s=0} &= U_o^2/\ell(1 + KU_o^2)
\end{aligned} \tag{13}$$

Here we see that steering sensitivity is a function of speed, wheelbase ($a + b = \ell$), and the stability factor, K. At low speeds the car follows a path whose curvature (C) is proportional to wheel deflection and inversely proportional to wheelbase (i.e., Ackermann steering):

$$C = \delta_w/\ell$$

Then the heading rate due to following this curved path is proportional to velocity

$$r = CU_o = U_o\delta_w/\ell$$

Therefore,

$$\frac{r}{\delta_w} = U_o/\ell$$

At higher speeds the stability factor exerts additional speed effects (Equation 13), with understeering cars ($K > 0$) having less speed sensitivity and oversteering cars ($K < 0$) having increased speed sensitivity.

Driver Behavior

Given the above approximate lateral response dynamics let us now analyze the corresponding driver control dynamics for the Y_y and Y_ψ blocks of Figure 1. The Y_ψ block includes several components of driver behavior. First there is a component due to basic limitations in the driver's response properties. These limitations can be subdivided further into subcomponents: pure time delays due to central nervous system processing and neural conduction to the limbs; and interface effects due to the spring mass damping system formed by the driver's arms coupled to the steering system. Most of the neuromuscular effects are high frequency and can be approximated by a pure time delay, $e^{-\tau s}$, in the frequency range of interest for car control (Reference 4).

The second Y_ψ component is a lead or anticipation term ($T_L s + 1$) that the driver adopts to counteract vehicle response characteristics discussed above. The third component is a gain K_ψ which sets the magnitude of δ_w corrections for given heading errors (ψ_e). Combining the above components we derive a heading response function that has been developed and used in a variety of past studies (e.g., References 2 and 5):

$$Y_\psi = K_\psi(T_L s + 1)e^{-\tau s} \tag{14}$$

A pure gain feedback for lane position errors has been found satisfactory in previous research:

$$Y_y = K_y \tag{15}$$

The addition of weighted lane position error and heading angle error can actually be considered as a composite angular error as shown in Figure 1:

$$\psi_L = K_y y_e + \psi_e \tag{16}$$

It has been previously shown that this equivalent angular error can be interpreted as the angular error to a projected aim point located some distance down the road, as shown in Figure 3 (Reference 2). The aim point concept is appealing because it represents somewhat of a perceptual efficiency for the driver. Instead of separately

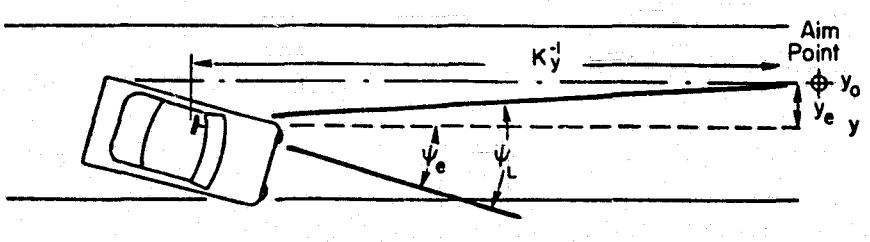


Figure 3. Aim Point Control Law: $\psi_L = \psi_e + y_e K_y$ where $\tan^{-1} y_e K_y \cong y_e K_y$

perceiving both lane position and heading errors, he need only perceive the angular error ψ_L to the aim point down the road.

One additional term will be added here that has received only scant attention in the literature (References 6 and 7). Analysis and modeling of driver response data at STI has shown the need for low-frequency characteristics or compensation that will act to reduce lane position error offsets and speed up the driver model's transient response and error reduction in tasks such as lane changes. The need for low frequency compensation can be seen by considering the effect of constant input disturbances, δ_d, to the Figure 1 block diagram. Given a constant δ_d input the driver model (i.e., Y_y and Y_ψ) discussed so far would have to allow constant lane position errors in order to generate a compensating wheel angle δ_w at the differential summing block. In the real world this would mean that drivers subjected to steady crosswinds or crowned roadways (i.e., inputs causing a constant force input to the vehicle) would drive with a constant lane position error.

The above offset error effect is not very reasonable, and the model can be corrected to eliminate it by adding a parallel trimming integrator, as illustrated in Figure 4, somewhere in the driver model feedforward path. The effect of the parallel integrator is to continue increasing its output in the face of steady errors, and holding its value as the errors approach zero. This will then produce steady wheel angles, δ_w, to compensate for steady disturbances, δ_d, without requiring a steady offset error.

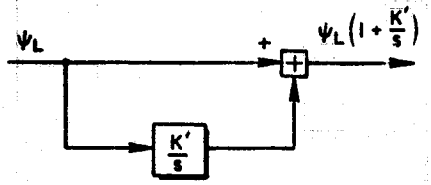


Figure 4. Parallel or "Trim Integrator for Counteracting Steady-State Error

There are two possible locations for the parallel integrator. One is in the Y_y block that would operate only on lane position errors. Another possibility is the Y_ψ block, where the parallel integrator would operate on the composite ψ_L signal which is a combination of lane position and heading errors. The Y_ψ location seems more reasonable for two reasons. First, the concept of perceiving a simple aim point error ψ_L would still be valid, which would not be true if the parallel integrator were applied to just the lane position errors. Second, steady-state heading errors can also develop in situations such as following curved paths, and the Y_ψ parallel integrator location would tend to compensate for these errors quicker than waiting for the heading error to accumulate into lane position errors which are operated on by the Y_y block.

Although we have rationalized the inner-loop (Y_ψ block) as the best location for the parallel integrator characteristic, we will analyze both possible locations (Y_ψ and Y_y) below in order to accumulate further evidence for the best location. Summarizing the driver response behavior for both parallel integrator locations we have:

Inner-Loop (Y_ψ) Parallel Integrator:

$$Y_\psi = \frac{s + K'}{s} K_\psi(T_L s + 1)e^{-\tau s}$$

$$K_y = K_y \qquad (16)$$

Outer-Loop (Y_y) Parallel Integrator:

$$Y_\psi = K_\psi(T_L s + 1)e^{-\tau s}$$

$$Y_y = \frac{s + K'}{s} K_y \qquad (17)$$

Driver/Vehicle Dynamics

Given the above dynamic characteristics for the vehicle and driver, let us now analyze the overall driver/vehicle system response given by Equation 5. The first term in Equation 5 (Y_ψ) is a driver characteristic, and the third is the car heading response (K_s is simply the steering ratio). The last term in Equation 5 is a combination of driver and vehicle characteristics. The interaction of these characteristics when K_y is a pure gain has been considered previously (Reference 2). Here we will consider the case where the parallel integrator is in the outer loop.

Combining Equations 5, 12, and 17 for the Y parallel integrator we end up with the following expression for the bracketed term in Equation 5

$$Y_y \frac{G_\delta^y}{G_\delta^\psi} + 1 = \frac{bK_y(s + K')[s^2 + (s/T_r) + (U_o/bT_r)]}{s^2(s + T_r^{-1})} \tag{18}$$

This can be seen to be a classical root locus problem, with one first-order zero (at K'), a second-order zero pair, two poles at the origin and a pole due to the heading time constant ($s = -1/T_r$). We can now vary the lane position gain (K_y) and plot the locus of roots for Equation 18. Root locus plots for vehicle characteristics used in past research (Reference 8) are compared in Figure 5. Here we see that with the parallel integrator in the outer loop a complex pair of low-frequency zeros occurs at reasonable values of K_y.

With the parallel integrator in the inner loop (Y_ψ, Equation 16) the result is two real zeros, one at K' and the other closed-loop zero due to the Equation 18 expression without the parallel integrator term. However, as illustrated in Figure 5, we see that for a large heading time constant the outer-loop parallel integrator always results in a complex zero with fairly low damping. This implies that the driver's low-frequency phase curve has a steep slope. However, data considered below will show that the low-frequency phase curves are never very steep and can only be fitted with two real roots as opposed to the Figure 5 complex roots.

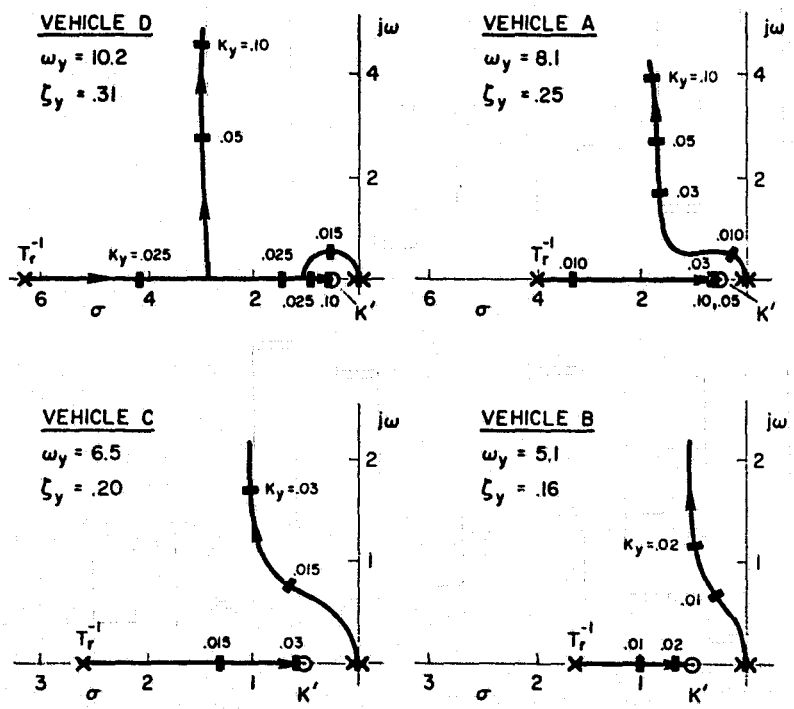


Figure 5. Root Locus for Outer-Loop Parallel Integrator (Vehicle dynamics described in Figure 6)

Phase Margin Approximation

The vehicle heading response (G_δ^ψ) dictates the high-frequency lead compensation to be provided by the driver. Plots of G_δ^ψ for various levels of vehicle heading time constant (T_R) studied in previous research (Reference 8) are given in Figure 6. The numerator zero and second-order denominator combine to give a net first-order-appearing transfer function combined with the kinematic integration, which gives an

VEHICLE CONFIGURATION	STEERING RATIO, K_s^{-1}	UNDERSTEER/OVERSTEER GRADIENT, K		INVERSE HEADING TIME CONSTANT, $T_r^{-1}(sec^{-1})$	INVERSE EQUIVALENT TIME CONSTANT, $T_{eq}^{-1}(sec^{-1})$	HEADING RESPONSE PARAMETERS	
		$10^4 \times sec^2/ft^2$	deg/g			ω_1 (rad/sec)	ζ_1
A ———	25:1	1.1	1.9	4.0	4.9	4.5	0.79
B —·—	17:1	1.3	2.2	1.6	2.3	2.0	0.77
C —··—	9:1	3.3	5.7	2.6	3.9	3.5	0.61
D -----	11:1	3.4	5.8	6.4	7.2	6.9	0.65

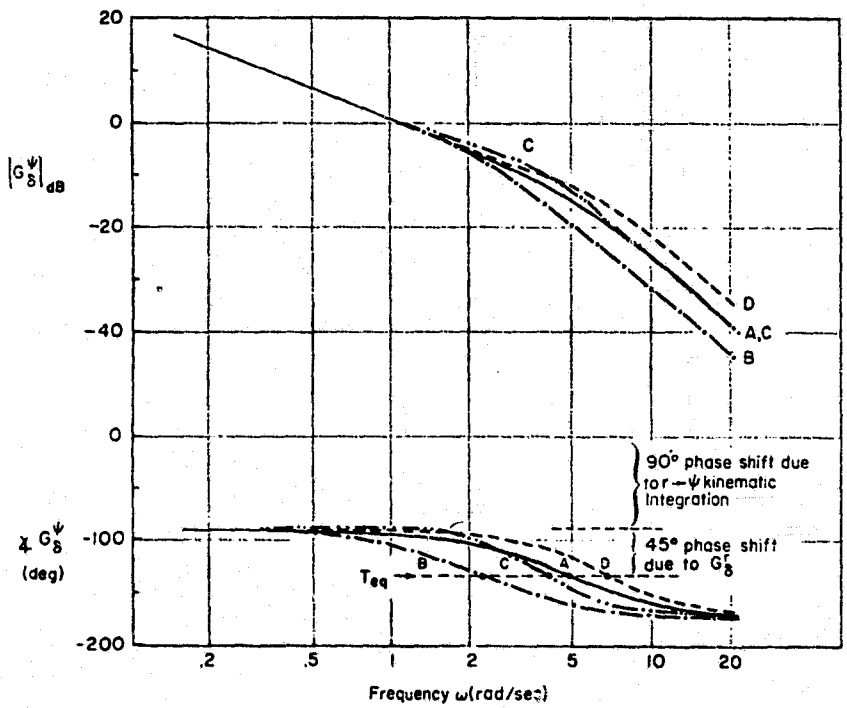


Figure 6. Test Vehicle Heading Response Transfer Functions

additional free s in the denominator. Inspection of the Figure 6 phase curves shows that the heading response dynamics can be described by an equivalent time constant, T_{eq}, which should dictate required driver lead compensation (i.e., $T_L \cong T_{eq}$).

Given the above components the composite driver/vehicle open-loop transfer function can be written as follows:

$$\frac{\delta_w}{\delta_e} = \overbrace{\frac{K_s K_\psi (s + K')(T_L s + 1)}{s}}^{\text{Driver Operation on Aim Point Error}} e^{-\tau s} \times \overbrace{\frac{N_\delta (s + T_r^{-1})}{s(s^2 + 2\zeta_1\omega_1 s + \omega_1^2)}}^{\text{Vehicle Heading Response}}$$

$$\times \underbrace{\frac{[s + (T_r')^{-1}](s + \alpha)}{s(s + T_r^{-1})}}_{\substack{\text{Driver/Vehicle Interaction} \\ \text{Due to Outer-Loop Closure}}} \qquad (19)$$

where $\alpha = K_y U_o$. This equation can be rearranged according to frequency characteristics:

$$\frac{\delta_w}{\delta_e} = \overbrace{\frac{(s + K')(s + \alpha)}{s^2}}^{\substack{\textcircled{1} \\ \text{Low-Frequency Compensation}}} \times \overbrace{\frac{K_\psi K_s N_\delta}{s}}^{\substack{\textcircled{2} \\ \text{Low Frequency K/s Slope}}} \times \overbrace{\frac{(s + T_r^{-1})}{(s^2 + 2\zeta_1\omega_1 s + \omega_1^2)}}^{\substack{\textcircled{3} \\ \text{Mid-High Frequency Vehicle Heading Response}}} \qquad (20)$$

$$\times \underbrace{\frac{s + (T_r')^{-1}}{s + T_r^{-1}}}_{\substack{\textcircled{4} \\ \text{Mid-High Frequency Interaction Due to Driver/Vehicle Interaction}}} \times \underbrace{(T_L s + 1)}_{\substack{\textcircled{5} \\ \text{Driver Lead Compensation for Vehicle Heading Response Lag}}} \times \underbrace{e^{-\tau s}}_{\substack{\textcircled{6} \\ \text{Driver High-Frequency Delay Limitation}}}$$

Now assume the driver adjusts T_L to compensate for combined phase lag characteristics of Terms ③ and ④. The 45 deg phase lag point of ③ is somewhat higher than T_r^{-1} and defines T_{eq}. Additional phase lead is derived from ④ so that the T_L^{-1} frequency break may be somewhat higher than T_r^{-1} or ω_1.

Given the complete transfer function expression above, it is now useful to consider an extended crossover model approximation. Assume that Expressions ③ and ④ combine into an equivalent first-order lag which is cancelled by the driver's lead term:

$$\frac{s + T_r^{-1}}{s^2 + 2\zeta_1\omega_1 s + \omega_1^2} \times \frac{s + (T_r')^{-1}}{s + T_r^{-1}} \times (T_L s + 1) \cong \frac{T_r^{-1}}{\omega_1^2} \qquad (21)$$

Then the driver/vehicle equivalent open-loop transfer function reduces to

$$\frac{\delta_w}{\delta_e} = \frac{(s + K')(s + \alpha)}{s^2} \times \frac{\omega_c}{s} \times e^{-\tau e s} \qquad (22)$$

where

$$\omega_c = \frac{K_\psi K_s N_\delta T_r^{-1}}{\omega_1^2} = \frac{K_\psi K_s U_o}{\ell}$$

for a neutral steering car and τ_e is an effective system time delay that accounts for the driver's time delay (τ) and residual phase lags (or lead) left over from the approximations in Equation 21.

Now we can analyze Equation 22 to determine the stability limit for the driver K_ψ gain. For stability the unity magnitude of Equation 22 must occur below the 180 deg phase lag point. At high frequencies ($\omega >> K'$ and α) the unity gain point is given by ω_c (i.e., the unity gain crossover frequency) and the Equation 22 phase is given by

$$\frac{\delta_w}{\delta_e} = -\frac{\pi}{2} - \tan^{-1}\frac{K'}{\omega_c} - \tan^{-1}\frac{\alpha}{\omega_c} - \tau_e\omega \qquad (23)$$

Phase margin is defined at the gain crossover frequency ω_c, which occurs at relatively high frequency, so that

$$\tan^{-1}\frac{K'}{\omega_c} \cong \frac{K'}{\omega_c} \quad ; \quad \tan^{-1}\frac{\alpha}{\omega_c} \cong \frac{\alpha}{\omega_c}$$

Thus the phase margin, or amount of extra phase shift allowable before instability is reached, is given by

$$\phi_M = 2\pi - \frac{\delta_w}{\delta_e}_{\omega_c} = \frac{\pi}{2} - \frac{K' + K_y U_o}{\omega_c} - \tau_e\omega_c \qquad (24)$$

The second term on the right side of Equation 24 is the phase lag due to the driver's low frequency behavior in controlling lane deviations. The third term is the phase lag due to the driver's basic time delay limitation, which defines the limiting bandwidth he/she can achieve. Note that the outer-loop operations (K' and K_y) add phase lag and further limit the achievable bandwidth, so that the driver must trade off inner- and outer-loop gains (i.e., K_ψ and K', K_y) in order to optimize performance.

Model Validation

A field study using an instrumented car on a closed test course has been previously reported on in Reference 8. The dynamics of the test vehicle could be easily modified, and the conditions given in Figure 6 were included in a test program involving 8 males and 8 females, ages 25-40, with an average of 13 years driving experience. Describing functions were obtained for each set of vehicle dynamics using a measurement technique reported previously (Reference 2).

Averaged driver describing function data for each set of vehicle dynamics are shown in Figure 7, along with curve fits according to the model of Equations 19 and 20. Some data reinterpretation over that reported in Reference 8 was necessary in order to obtain detailed model fits. Although each vehicle data set was fit individually, some constraints were observed across vehicles in order to obtain parameter values that changed in an orderly manner with vehicle heading time constant.

Model parameters are plotted as a function of inverse equivalent vehicle time constant (i.e., the basic bandwidth of vehicle heading or yawing response) in Figure 8. In general, the trends shown in Figure 8 are consistent with the known behavior of the human operator (Reference 4), to wit:

- Lead generation (T_L) increases with increasing system lag (T_{eq}), although the cancellation is not complete.
- Operator time delay (τ) increases with lead generation (T_L).
- Open-loop gain ($K_\psi K_c$) decreases with increasing system lag to maintain stability.

In this case of multiple-loop dynamics we also note that the outer-loop gain (K_y) decreases with increasing vehicle lag, which according to Equation 24 also tends to maintain stability.

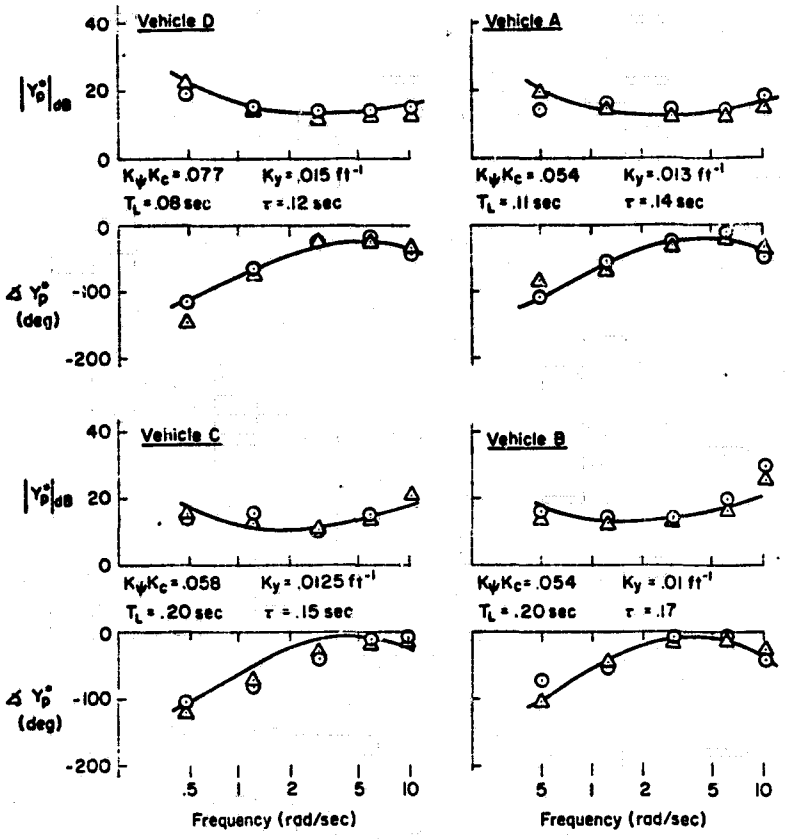


Figure 7. Field Test Driver/Vehicle Describing Functions and Model Fits (⊙ Male; △ Female)

The equivalent crossover model parameters plotted in Figure 9 give further insight into the driver's adaptation to different vehicle dynamics. Here we see that the phase lag component due to the driver's lane control behavior $[(K' + K_y U_o)/\omega_c]$ is maintained relatively constant. This is accomplished by reducing K_y as ω_c is reduced in response to increased system time delay (τ_e) which in turn results from increased vehicle lags.

CONCLUDING REMARKS

The approximate driver/vehicle steering dynamics analysis developed herein provides insight into the driver's adaptive behavior. The driver offsets increased vehicle lags with anticipation or lead behavior, but in doing so incurs additional time delay penalty. The driver then compensates for the phase lag due to extra time delay by reducing his/her gain.

The effect of the driver's inherent time delay penalty on stability is analyzed with a phase margin approximation. This approximation shows that both crossover frequency and outer-loop gain affect steering stability. Thus, the driver's behavior in controlling lane position results in a phase lag penalty which influences the directional stability of the driver/vehicle system.

The analysis in this paper relates to directional control stability independent of the path the driver is commanded to follow. Path commands due to roadway curvature evoke additional driver behavior which has been considered previously in References 3 and 5.

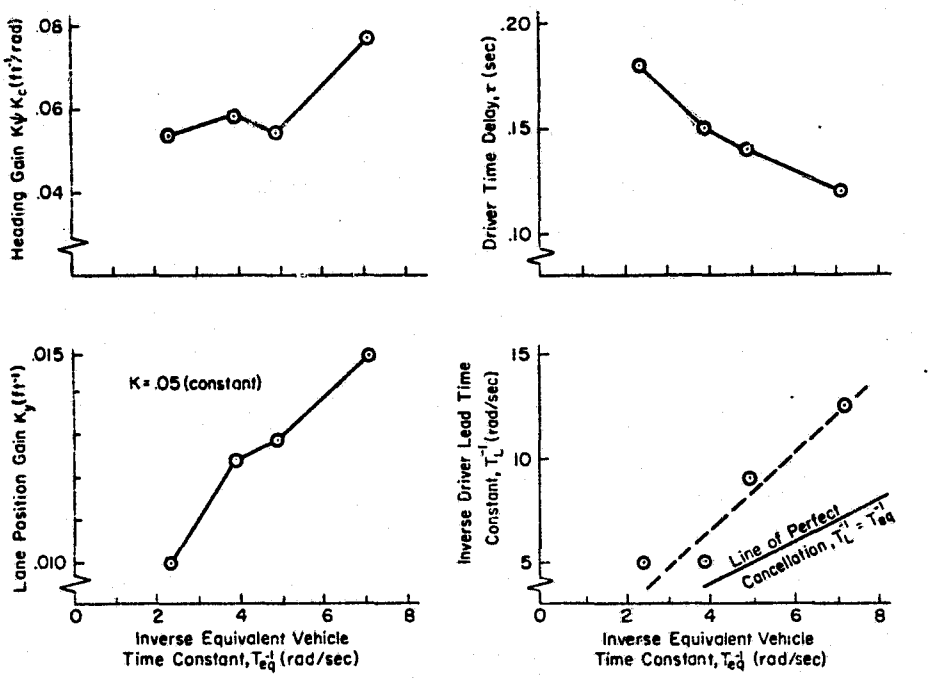


Figure 8. Effects of Vehicle Dynamics on Driver Model Parameters

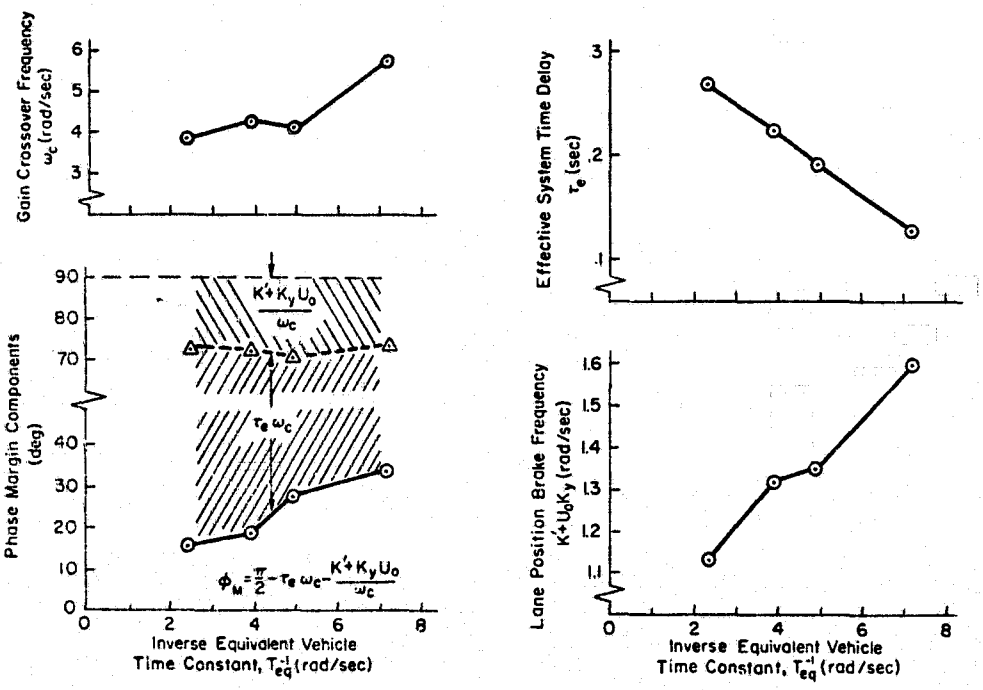


Figure 9. Effects of Vehicle Dynamics on Driver Crossover Model Parameters

REFERENCES

1. McRuer, D. T.: Simplified Automobile Steering Dynamics for Driver Control. Presented at SAE Aerospace Control and Guidance Systems Committee Meeting No. 35, Palo Alto, CA, 19-21 Mar. 1975 (Systems Technology, Inc., P-165).

2. McRuer, D. T.; Weir, D. H.; Jex, H. R.; Magdaleno, R. E.; and Allen, R. W.: Measurement of Driver/Vehicle Multiloop Properties with a Single Disturbance Input. In IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-5, No. 5, Sept. 1975, pp. 490-497.

3. Allen, R. W.; and McRuer, D. T.: The Effect of Adverse Visibility on Driver Steering Performance in an Automobile Simulator. In SAE Transactions, 1977, pp. 1081-1092.

4. McRuer, D. T.; and Krendel, E. S.: Mathematical Models of Human Pilot Behavior. AGARD-AG-188, Jan. 1974.

5. Allen, R. W.; and McRuer, D. T.: Driver Steering Dynamics Measured in a Car Simulator Under a Range of Visibility and Roadmarking Conditions. In Proceedings of the 13th Annual Conference on Manual Control. NASA CR-158107, 1977, pp. 180-196.

6. Delp, P.; Crossman, E. R. F. W.; and Szostak, H.: Estimation of Automobile Driver Describing Functions from Highway Tests Using the Double Steering Wheel. In Proceedings of the 7th Annual Conference on Manual Control. NASA SP-281, 1971, pp. 223-236.

7. Crossman, E. R. F. W.; and Szostak, H. T.: Man-machine Models for Car Steering. In Proceedings of the 4th Annual Conference on Manual Control. NASA SP-192, 1969, pp. 171-195.

8. McRuer, D.; and Klein, R.: Effects of Automobile Steering Characteristics on Driver/Vehicle Performance for Regulation Tasks. Paper 760778, presented at SAE Automobile Engineering Meeting, Detroit, 18-22 Oct. 1976.

D53

DETECTING HUMAN OPERATOR IMPAIRMENT WITH A PSYCHOMOTOR TASK*

By R. Wade Allen, Anthony C. Stein, and Henry R. Jex

Systems Technology, Inc.
Hawthorne, California 90250

SUMMARY

Psychomotor tasks have long been used as instruments for measuring human operator behavior and have often been proposed as selection and training devices. This paper discusses the application of one such test, the "Critical-instability Tracking Test" (CTT, often called the Critical Task), in the detection of alcohol impaired human operators.

The validity of the CTT as a measure of human operator performance and its sensitivity to alcohol impairment have previously been established. The current work to be presented here concerns the detection problem wherein statistical decision theory is used to optimize the test's ability to discriminate impaired operators in the face of inherent variability in human performance.

Testing strategy is discussed including setting performance criteria levels and determining the number of attempts permitted in order to pass the test. Also, the task's ability to detect operator impairment is compared to performance measures obtained in a driving simulator.

INTRODUCTION

The development of methods and devices for measuring visual motor behavior has been of continuing interest for decades. The pursuit rotor (Reference 1) is a classic example of a psychomotor task that has been used in numerous experiments over the years. During the second world war there was interest in the use of psychomotor tasks for personnel selection (Reference 2). More recently, with increasing sophistication in the measurement of manual control behavior, psychomotor tracking

---

*The work reported herein is part of an ongoing research program sponsored by the National Highway Traffic Safety Administration (Field Tests of the Drunk Driving Warning System, DOT-HS-8-02052). The contents of this paper represent the views of the authors and do not necessarily reflect the official views or policy of the Department of Transportation.

tasks have been used in applications such as monitoring the progress in treatment of neurological disorders (Reference 3).

Operator visual-motor behavior is an important skill in the control of modern man-machine systems. Efficient psychomotor testing procedures could provide a means for operator selection, proficiency checks and screening for impaired functioning due to fatigue, stress, alcohol, drugs, etc. To be appropriate for such applications the psychomotor task must measure behavior that is both relevant to performance of the real world task and efficient in its administration. The relevance and convenience of a psychomotor task in a given application must be decided on a case by case basis. The measurement efficiency or test power is a more general concept, however, and is the primary concern of this paper.

The test power of a psychmotor task depends on the statistical properties of the performance measure and the magnitude of change it is desireable to detect. In this paper we will consider a general statistical model for psychomotor performance measures. This statistical model will be applied to a well developed and tested psychomotor test, the Critical Tracking Test (Reference 4). We will analyze the task's ability to detect alcohol impaired operators, and show how multiple trial strategies can be used to optimize the task's discriminability and efficiency. In the next section we will present a statistical model and interpret past results with the Critical-instability Tracking Test (CTT) in the light of this model.

STATISTICAL MODEL

Components

In order to obtain stable estimates of a performance metric, multiple trials are usually taken. Various strategies can then be utilized, as will be discussed in the next section, to reduce the effects of variability in the measure. In order to evaluate the effect of various strategies, a statistical model is useful. A general model will be considered here that partitions a performance metric into a series of deterministic and random components:

Net Score		Population Average		Stress Impairment	Trial Order	Long Term Trends		
λ_c	=	λ_{c_o}	+	$\Delta\lambda_1$	+ $\Delta\lambda_2$ +	$\Delta\lambda_3$	...	Deterministic Effects

(1)

	Within a Subject		Between Subjects	Subjects by Stress Interaction		
... +	ε_1	+	ε_2	+ ε_3		Random Effects

The deterministic effects describe what happens to the performance metric on the average. Here we have allowed for various effects:

- Stress applied to the human operator which would typically degrade performance.
- Trial order or short term trends having to do with warmup and/or fatigue effects causing a trend across a few repeated trials.
- Long term trends due to learning with repeated task experience.

Random effects concern statistical variations in performance, without any apparent causal factors:

- Within a subject - random variations across multiple trials which are over and above general trial-to-trial trends.
- Between subjects - random average performance differences between randomly selected subjects.
- Subject by stress interaction - random differences in response to a given stress between randomly selected subjects.

Equation 1 can be considered an Analysis of Variance (ANOV) model, and estimates of the various model components can be obtained using ANOV and regression analysis procedures on a given data base.

Critical Tracking Test Data

The Critical Tracking Test (CTT) has been used in a variety of research studies over more than a decade and good estimates of various Equation 1 terms are available. In Figure 1 for example, we see the effect of alcohol on CTT score decrements compared across several experiments conducted by different research groups with different experimental procedures (References 5-8). Note that CTT performance universally degrades by about 10 percent at the common legal limit of blood alcohol concentration (BAC) specified in many state vehicle codes, and more than twice this decrement at BAC's above 0.20 characteristic of heavy drinking.

Short and long term trends in CTT scores have been analyzed in several experiments. In Reference 9 the psychomotor performance of four subjects was analyzed who were isolated in a space station simulator for three months. CTT scores were obtained two times a day for 12 weeks. ANOV procedures did not show any day-to-day trends averaged across weeks, but did show a week-to-week improvement trend over the twelve weeks. We have also reanalyzed some other more recent CTT data bases (References 5-7). In this reanalysis we allowed for a trial-to-trial component for six trial blocks. The analysis showed negligible trial-to-trial effects.

The random variance components in Equation 1 can be estimated with ANOV procedures. This has been done for a wide variety of CTT experiments, and the results are summarized in Table 1. Due to equipment and training variations, the mean performance level varied considerably between some experiments. To minimize this factor normalized variability components are also given in Table 1, and it is the normalized levels that are noted to be quite consistent across various experiments. One standard deviation of within-subject variability runs about 9 percent (range of 6-10 percent), with between subject variability running about the same level. Subject-by-stress interaction runs somewhat less at about 5 percent on the average (range of 3-6 percent).

The Table 1 results tell us several things about setting up a decision making strategy using the Critical Tracking Test. First of all, within subject variability is about the same magnitude as the degradation due to legally drunk blood alcohol levels. Thus, detecting this magnitude of change in performance level will definitely require multiple samples for reasonable reliability. Secondly, the between -- subject variability component represents a significant portion of the total variability so that subjects should be used as their own control. Finally, after accounting for within and between subject variability components, a small amount of residual variability between subjects is still to be expected in their differential responses to a given stress.

In order to develop a statistical model for Critical Tracking Test scores, the Dunlap data bases summarized in Table 1 were subjected to a detailed reanalysis. Individual Score decrements were obtained by removing each subject's mean sober score from all his/her individual scores. The ensemble distributions of the resulting differential scores at four BAC levels for two experiments are shown in Figure 2. The cumulative Gaussian distribution scaling on Figure 2 (standard "probit" plots) helps reveal several characteristics of the CTT score distributions. First, the scores at each BAC are roughly Gaussianly distributed between the 5 percent and 95 percent points, as indicated by the straight slopes. Second, the variability of population decrements increases slightly with BAC (shallower slope). Since we have accounted for the between-subject variability by analyzing differential scores, the increasing variability with BAC amounts to a combination of trial-to-trial and subject-by-stress interaction variability discussed previously.

Based on analysis of the three Dunlap CTT data bases, we have developed the simple statistical model for alcohol effects on CTT single trial score means and variability shown in Figure 3. Assuming a Gaussian distribution, the analytical model of Figure 3 can now be used to analyze, synthesize and interpret the decision strategies discussed next.

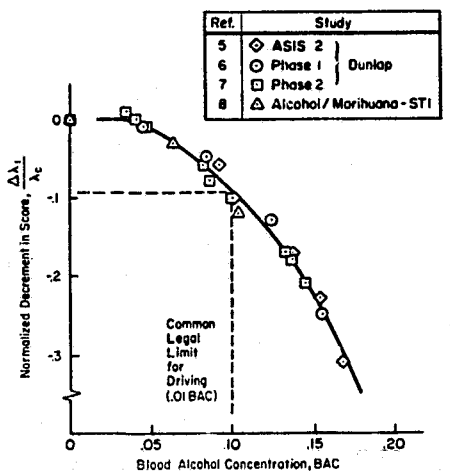


Figure 1. Alcohol Effects on Critical Tracking Test Scores

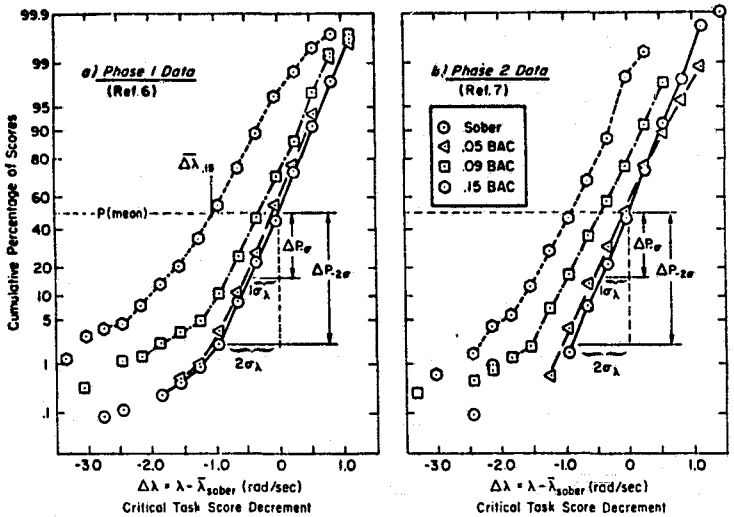
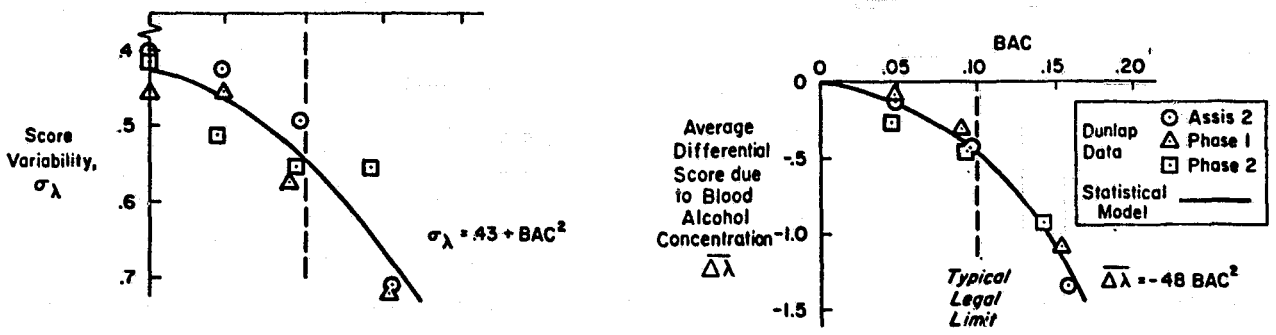


Figure 2. CTT Score Distributions Relative to Each Subject's Sober Mean

TABLE 1. VARIANCE COMPONENTS FOR CRITICAL TRACKING TASK DATA

Experiment		No. Of Ss	Group Avg $\bar{\lambda}_c$	Absolute Variability, σ			Normalized Variability, $\sigma/\bar{\lambda}_c$			Ref. No.
				Within Ss σ_{ϵ_1}	Between Ss σ_{ϵ_2}	Subject X Stress σ_{ϵ_3}	Within Ss $\sigma_{\epsilon_1}/\bar{\lambda}$	Between Ss $\sigma_{\epsilon_2}/\bar{\lambda}$	Subject X Stress $\sigma_{\epsilon_3}/\bar{\lambda}$	
ALCOHOL STRESS										
DUNLAP EXPERIMENTS	ASIS2	15	4.75	.421	.437	.293	8.9%	9.2%	6.2%	5
	Phase 1	24	4.64	.434	.476	.263	9.4%	10.3%	5.7%	6
	Phase 2	24	4.62	.442	.431	.197	9.6%	9.3%	4.3%	7
STI	Alcohol and Marijuana Tests	12	4.7	.32	.33	.28	6.8%	7.0%	6.0%	8
OTHER STRESSES										
	Broadband Noise	9	3.1	.28	.18	.12	9.1%	5.8%	3.9%	10
	90 Day Confinement	4	6.5	.41	.319	.333	6.2%	4.9%	5.1%	9
	Motion Effects	8	5.14	.514	.717	.134	10.0%	14.0%	2.7%	11



a) Differential Mean b) Standard Deviation

Figure 3. Statistical Model Parameters for Alcohol Effects on CTT Score

IMPAIRMENT DETECTION STRATEGY

Detection Criteria and Errors

Given a fair amount of knowledge about the statistical properties of CTT scores, we now would like to set up a decision strategy that will allow us to detect alcohol impaired psychomotor behavior. Given that an individuals' trained, sober performance capability is known, we wish to be able to detect when peformance is reliably below this criterion level. This presents a signal-in-noise detection problem; i.e., to detect the mean shift in score in the face of its instrinsic variability. This will require a multiple trial strategy since, as mentioned previously, the mean shift in score due to alcohol impairment to be detected is roughly the same magnitude as the standard deviation in CTT scores.

In the detection problem one wishes to minimize two types of classical statistical decision making errors. The first, referred to as a "Type I" error, (usually designated by the symbol α), is the probability that there is no difference in the mean performance (i.e., rejecting the null hypothesis) when the test outcome says there is a decrement. In the current application this would amount to a sober and unimpaired subject being rejected by the test sequence. The second, or "Type II" error (β) is the probability that there is a decrement (hence, impairment) when the test concludes there is none. Here, β would be the probability that an alcohol- or drug-impaired subject could pass the test sequence.

The approach taken in setting error levels is similar to the general experimental design problem, in that we will assign a tolerable level

for α, setting it at a relatively low value so that the unimpaired operator is not inconvenienced (i.e., can routinely pass the test). However, because test time is costly, one cannot simply take an arbitrary number of samples in order to achieve a low β level for reliably detecting a specified change in performance $\Delta\lambda$. In order to have a practical psychomotor screening test, one has to pick an efficient multiple trial strategy and attempt to achieve a low level of β at reasonable $\Delta\lambda$ (impairment) levels with a relatively few number of trials on the average, all subject to an acceptably low α.

In general, test score pass levels, λ_p are set some multiple K_σ of the test variance below the individual's sober baseline, so as to achieve a specified α based on Gaussian statistical assumptions (or from probit plots such as those in Figure 2):

$$K_{\sigma_n} = \frac{\bar{\lambda}_{sober} - \lambda_{pass}}{\sigma_{\lambda_n}} \qquad (2a)$$

This can be solved for the pass criterion:

$$\text{pass score: } (\lambda_p) = \bar{\lambda}_{sober} - K_{\sigma_n}\sigma_\lambda \qquad (2b)$$

where

λ_p = criterion level of λ_c above which a trial is passed

$\bar{\lambda}_{sober}$ = a given subjects trained, unimpaired performance level

σ_λ = standard deviation of scores for the test strategy used

K_{σ_n} = constant depending on the test strategy and number of trials, n; and desired level of α

For example, looking back at Figure 2b, for $\alpha \triangleq 0.025$ (sober fail probability at 2.5 percent) the single trial score decrement for sober tests (BAC=0) is about $(\Delta\lambda \triangleq) 2\sigma_\lambda$ below the sober mean; hence, $K_{\sigma_1} \triangleq 2.0$.

Thus the score decrement which could be reliably detected in a single trial ($\alpha = 0.025$) is $\Delta\lambda = -2\sigma_\lambda \triangleq -0.86$ rad/sec. This is much larger than the desired BAC = 0.10 decrement of about -0.48, so multiple trials will be required to increase test sensitivity.

Multi-Trial Test Strategy

The method of setting K_{σ_n} depends on the particular strategy employed. To illustrate the process by which various strategies give

decision statistics, the "test-outcome probability models" for two strategies are given in Figure 4, and the relevant statistics for three strategies are summarized in Table 2. Each of the strategies is described below.

Average of n trials strategy (Figure 4a). Averaging across multiple trials is a familiar approach (i.e., Reference 12) which improves detection because the standard deviation of the distribution of means is reduced over the single trial standard deviation by a factor of $\sqrt{n}$:

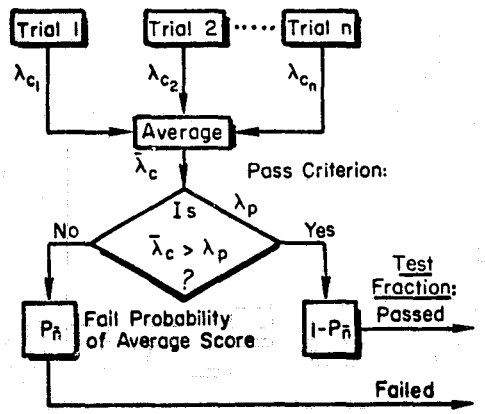
$$\sigma_{\bar{\lambda}_n} = \sigma_\lambda \; / \; \sqrt{n} \qquad (3)$$

Since the standard deviation of the mean is reduced over the single trial value, the pass criterion coefficient in Equation 3 is also similarly reduced

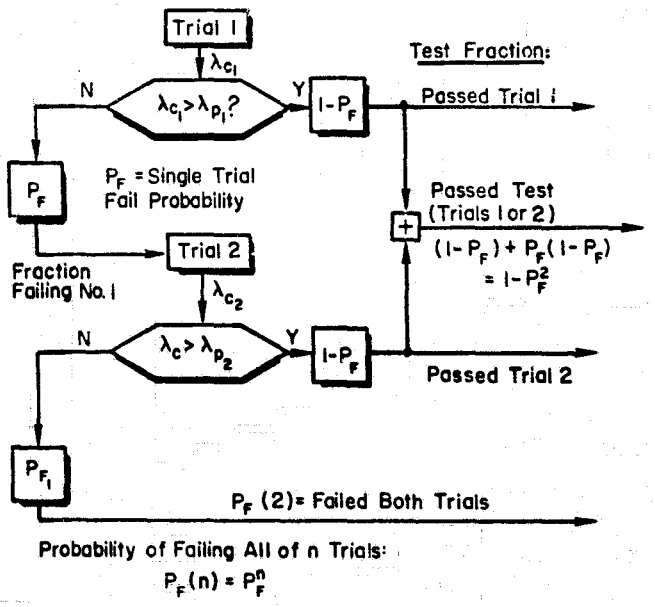
$$K_{\sigma_n} = K_{\sigma_1} \; / \; \sqrt{n} \qquad (4)$$

where K_{σ_n} is the single trial value of K to give a given α for the entire test.

One-of-n trials strategy (Figure 4b). In this strategy, the operator has n trials in which to attempt to exceed the pass criterion. When



a) Model For Averaging of n Trials



b) Model For Passing Any One of n Trials

Figure 4. Test-Outcome Probability Models

TABLE 2. DECISION STRATEGY SUMMARY

	STRATEGY		
	Average of n Trials	One of n Trials	Sequential
Decision Criteria	Pass: $\bar{\lambda} > \lambda_p$ Fail: $\bar{\lambda} \leq \lambda_p$	for i=1,.....n Pass: $\lambda_i > \lambda_p$ Fail: $\lambda_i \leq \lambda_p$ Take another trial: $i < n$ and $\lambda_i \leq \lambda_p$	for i=1,.....n Pass: $\lambda_i > \lambda_p$ Fail: $\lambda_i < \lambda_f$ Take another trial: $\lambda_f \leq \lambda_i \leq \lambda_p$
Type I error, α (unimpaired failure)	Prob. $[\bar{\lambda}_{sober} \leq \lambda_p]$ λ_{sober}: $N[\bar{\lambda}_{sober}, \sigma_\lambda/\sqrt{n}]$	P_F^n	$\frac{P_F}{P_P + P_F}$
Average number of unimpaired trials	n	$\frac{(1 - P_F)\sum_{i=1}^{n} n P_F^{n-1}}{1 - \alpha}$	$\frac{1 - P_F}{P_P}$

where P_F = single trial probability of failure
P_P = single trial probability of passing

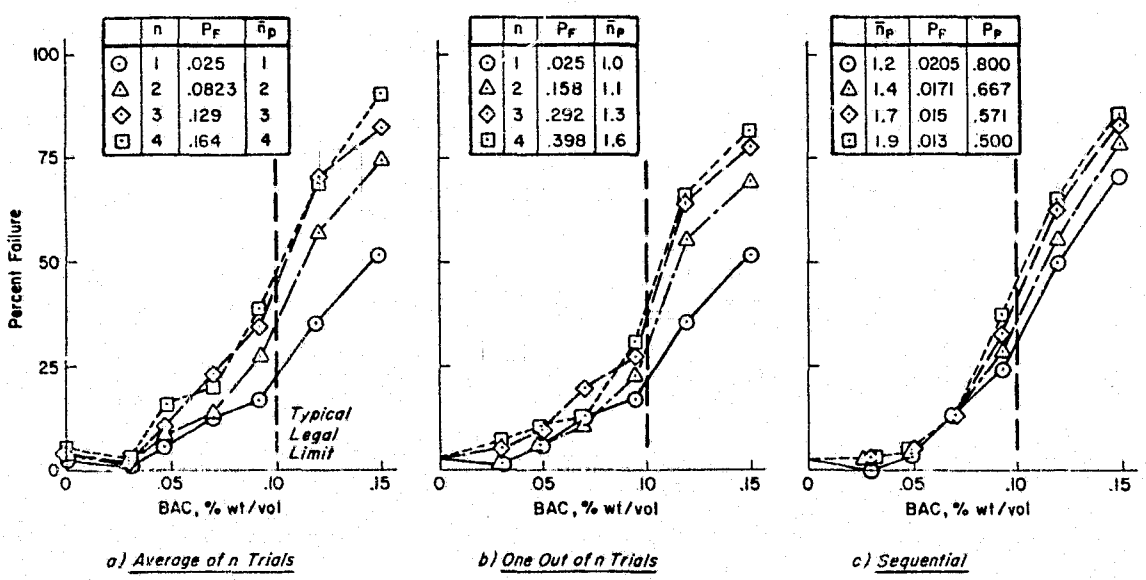


Figure 5. CTT Alcohol Impairment Discriminability for Various Decision Strategies Applied to Data in Refs. 6 and 7

the pass criterion is exceeded the test is passed, and no further trials are required. To fail the test, the operator must fail n trials in a row. If we assume that these are independent trials based on past data analysis, then the probability of failing the test is the product of all n of the single-trial failure probabilities:

$$P_F(n) = (P_{F_1})^n \tag{5}$$

Thus, for a specified level of α, the one trial probability of an unimpaired failure is given by:

$$P_{F_1} = \alpha^{1/n} \tag{6}$$

Given P_{F_1} we can then refer to Figure 2, or use tables for the cumulative Gaussian distribution function, to determine K_{σ_n}.

Sequential Strategy. This strategy has two criterion levels, an upper one for passing and a lower one for failing. If the operator scores in between these two levels, he/she is allowed to take another trial. This is a more general form of the one-of-n strategy, wherein the lower fail criterion is added so that the operator can pass or fail the test on an early trial. With this strategy one specifies α and the desired number of unimpaired trials for a decision, then uses the formulas in Table 2 to compute the single trial probabilities of passing (P_P) and failing (P_F) the test. The Table 2 formulas can be derived by noting that the probability of passing within n trials has a negative binomial distribution. Given P_P and P_F one can then use Figure 2 or Gaussian distribution tables to compute single-trial pass and fail scores.

$$\text{Pass Level:} \quad \lambda_P = \bar{\lambda}_{sober} - k_p\sigma_\lambda \tag{7a}$$

$$\text{Fail Level:} \quad \lambda_F = \bar{\lambda}_{sober} - k_f\sigma_\lambda \tag{7b}$$

More sophisticated sequential strategies are discussed in Reference 15.

Comparison of Strategies

The above strategies were applied to the extant CTT data bases (References 5-7) and the derived test failure rates as a function of related alcohol concentration are plotted in Figure 5. These curves show the discriminability (percent failing the test at given BAC levels) for various decision strategies, this is an index of test "benefit." Also tabulated on Figure 5 are the average number of trials required to pass the test when unimpaired ($\bar{n}_p$), which is an index of test "cost". In general, significant detection of impairment does not occur until BAC is in

the region of 0.10 and above. Comparing the three strategies, there is surprisingly little difference in impairment discriminability, vs. BAC. Also note that as n increases, the average number of trials for a sober pass (n_p) increases until a point of diminishing returns is reached relatively rapidly*.

The average-of-n trials strategy has an practical disadvantage, in that n trials are always required. For comparable discriminability, the other two strategies require less than half the average number of trials for a sober pass. The sequential strategy has a disadvantage in that the one trial probability of failure (P_F) levels are quite low and occur out in the tail of the score distribution (see Figure 2). As noted in Figure 2, this portion of the distribution is not very reliable and is probably subject to low score outliers. Achieving a reliable value for α would be difficult in general with the sequential strategy. In contrast, the single trial failure probabilities for the one-of-n strategy place the pass criteria more towards the center of the distribution, as noted earlier, which leads to relative reliability in setting α.

Based on the above considerations, the one-of-n strategy appears to be an optimum test strategy for detecting impairment with the CTT. This strategy also has a subjective appeal in that the unimpaired subject usually passes on the first or second trial, and is not penalized for one individually poor trial (which could arise due to distractions, getting a bad start, etc.). The impaired subject is given a maximum number of chances to pass, which is both fair and (desirably) inconvenient. The next section describes an experiment designed to validate the one-of-n strategy against other relevant measures of psychomotor impairment.

VALIDATION EXPERIMENT

Procedures

To validate the effectiveness of the CTT and test strategy just described, an experiment was conducted that compared CTT score with both BAC and driving performance in the STI Driving Simulator (Reference 13). Subjects were convicted drunk drivers obtained through the cooperation of the Los Angeles Municipal Courts. Twenty-four so called volunteers were permitted to participate in the experiment as a condition of probation, and, in exchange, received a reduction on their fine.

Once accepted, subject's were required to participate in three 2-hour training sessions and three full-day experimental sessions. A complete discussion of the training problem is found in a companion paper (Reference 14). Each subject participated in one placebo and two

---

*This trend is consistent with so called "operational characteristic curves" used in classical experimental design, Reference 12.

drinking sessions. The subject population was divided into three groups matched for age, sex, and driving experience, and the order of occurrence of the placebo session was different for each group.

Results

The discriminability results agree with the statistical model, as Figure 6 shows. At the lower BAC levels the actual tests were slightly more sensitive than the model predicts. This result validates the statistical model and ties in this data base with pass experiments.

While the above results prove that the one-of-n strategy CTT Failure Rate is a reliable correlate of BAC, the more important data is the correlation between driving performance, and both CTT failure rate and BAC. Figure 7 makes this comparison for both accidents and speed exceedances. In both figures, CTT Failure Rate (one-of-four pass strategy) is shown by the solid hexagonal symbol, and simulator driving performance by the open symbols. As CTT Failure rate increases, the actual driving performance suffers at almost the same rate implying an excellent correlation between CTT Test failure rate and simulator driving impairment*.

Having noted in Figure 7 that the accident detection rate is similar to the basic test failure rate, we assessed whether test failure discriminated against simulator accidents and speed exceedances. These comparisons, shown in Figure 8, indicate that at 0.15 BAC, a pre-drive CTT failure detected 81 percent of the subsequent accidents. This accident detection rate at 0.15 BAC is comparable to the CTT failure rate shown in Figure 6.

These correlations between predicted and actual test performance show that it is now possible to both predict and verify vehicle operator impairment using a psychomotor task such as the Critical Tracking Test.

CONCLUDING REMARKS

The procedures discussed here for evaluating and optimizing the discriminability of psychomotor tests such as the CTT are fairly general and should be applicable to other visual-motor tasks as well. The application discussed here involved the detection of operator impairment which requires individualized criterion levels for each subject. These procedures would also be applicable to population screening and selection applications as well, where a single population criterion level

---

*It has been shown previously (Ref. 13) that the STI driving simulator accident rate vs. BAC corresponds with large scale epidemiological data bases.

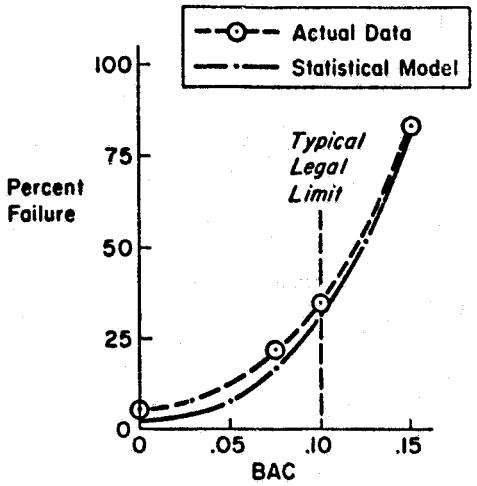


Figure 6. CTT Alcohol Impairment Discriminability Validation Data Compared with Statistical Model Prediction

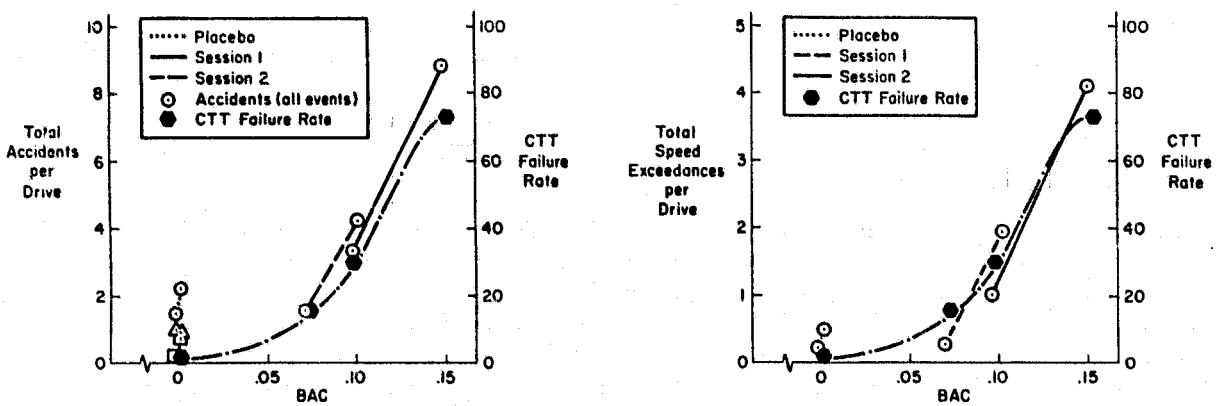


Figure 7. Comparison of Driving Simulator Performance CTT Failure Rate

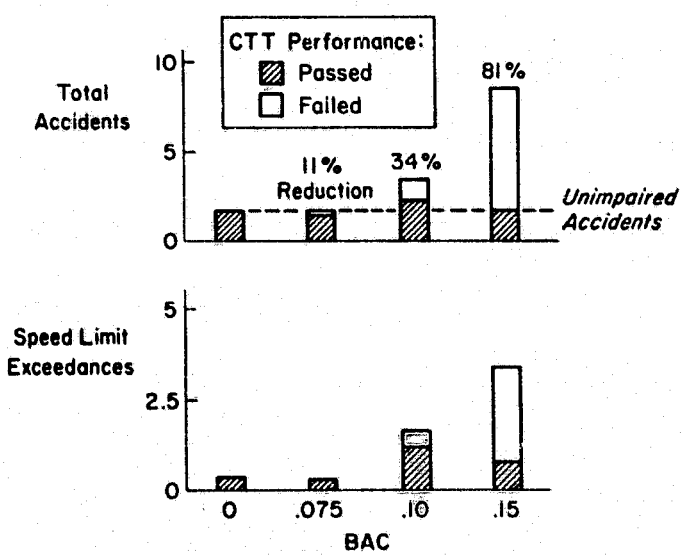


Figure 8. CTT Discriminability of Impaired Driving Simulator Performance

would be used that corresponded to desired or established performance levels. Future tasks to be used in similar applications should be evaluated according to these procedures, where the Type I error is specified and held constant while evaluating various multiple trial strategies with respect to their discriminability.

The evaluation procedures discussed herein have also illustrated some favorable features of the Critical Tracking Test for measuring human operator impairment. These include near-universal sensitivity of all members of the population, good test/retest reliability, excellent relevance to driving performance and reasonable training requirements. The test power and efficiency of other tasks should be evaluated in this light.

Relevance of a particular detection test to a given application is a continuing issue, that must be addressed by research designed to demonstrate a test's ability to detect the ultimate problem (e.g., accidents; rather than BAC) as illustrated by the validating experiment described here.

REFERENCES

1. Seashore, R. H., "Stanford Motor Skills Unit," Psychology Monographs, Vol. 39, 1928, pp. 51-66.

2. Stuit, D. B. (Ed.), Personnel Research and Test Development in the Bureau of Naval Personnel, Princeton University Press, 1947.

3. Repa, B. S., et al, "Clinical Applications of Tracking," in Proceedings of the Eighth Annual Conference on Manual Control, AFFDL-TR-72-92, May 1972, pp. 343-367.

4. Jex, H. R., J. D. McDonnell, and A. V. Phatak, A "Critical" Tracking Task for Man-Machine Research Related to the Operator's Effective Delay Time, Part I: Theory and Experiments with a First-Order Divergent Controlled Element, NASA CR-616, Nov. 1966.

5. Oates, John F., Jr., Experimental Evaluation of Second-Generation Alcohol Safety-Interlock Systems, Transportation Systems Center DOT-TSC-NHTSA-73-9, 1973.

6. Oates, John F., Jr., David F. Preusser, Richard D. Blomberg and Marlene S. Orban, Laboratory Testing of Alcohol Safety Interlock Systems, Vol. I: Procedures and Preliminary Analyses, Dunlap and Associates, Inc., Jun. 1975.

7. Oates, John F., Jr., David F. Preusser, and Richard D. Blomberg, Laboratory Testing of Alcohol Safety Interlock Systems, Phase II, Dunlap and Associates, Inc., Dec. 1975.

8. Allen, R. W., S. H. Schwartz, A. C. Stein, et al, The Effects of Alcohol and Marihuana on Driver Control Behavior, Vol. I: Laboratory Simulation Experiment, STI ITR 1066-1, Apr. 1978.

9. Allen, R. W., and Henry R. Jex, Visual-Motor Response of Crewmen During a Simulated 90-Day Space Mission as Measured by the Critical Task Battery, STI TR-1012-1, Nov. 1971.

10. Allen, R. Wade, Raymond E. Magdaleno, and Henry R. Jex, Effects of Wideband Auditory Noise on Manual Control Performance and Dynamic Response, AMRL-TR-75-65, Oct. 1975.

11. Jex, Henry R., Richard J. DiMarco, and Warren F. Clement, Effects of Simulation Surface Effect Ship Motions on Crew Habitability -- Phase II. Vol. 3: Visual-Motor Tasks and Subjective Evaluations, STI TR-1070-3, Feb. 1976.

12. Bowker, Albert H., and Gerald J. Lieberman, Engineering Statistics, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1959.

13. Allen, R. Wade, and H. R. Jex, "Driving Simulation -- Requirements, Mechanization and Application, SAE Paper No. 800448, presented at SAE Congress and Exposition, Cobo Hall, Detroit, Feb. 1980.

14. Cook, Marcia, Henry R. Jex, Anthony C. Stein, and R. Wade Allen, "Using Rewards and Penalties to Obtain Desired Subject Performance," 1981 Annual Conference on Manual Control Proceedings, Jun. 1981, in press.

15. Wald, A., Sequential Analysis, New York, J. Wiley & Sons, Inc. 1962.

VOICE CONTROL OF THE SPACE SHUTTLE VIDEO SYSTEM

A. K. Bejczy and R. S. Dotson
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

J. W. Brown and J. L. Lewis
Spacecraft Design Division
Johnson Space Center
Houston, TX 77058

SUMMARY

The feasibility and utility of controlling the Space Shuttle TV cameras and monitors by voice has been investigated. The voice control application concept is related to task scenarios where the operator uses both hands to control the 50-foot (16-meter) manipulator of the Space Shuttle. The use of computer-recognized voice commands allows the operator to effectively press the control buttons of the Shuttle TV cameras and monitors by voice while he manually controls the Shuttle manipulator. The pilot voice control system developed at the Jet Propulsion Laboratory (JPL) to test and evaluate the feasibility of controlling the Shuttle TV cameras and monitors by voice commands utilizes a commercially available discrete word speech recognizer which can be trained to the individual utterances of each operator. Successful ground tests have been conducted with this pilot application system at the Johnson Space Center (JSC) Manipulator Development Facility (MDF) using a simulated full-scale Space Shuttle manipulator. The test configuration involved the berthing, maneuvering and deploying a simulated science payload in the Shuttle bay. The handling task typically required 15 to 20 minutes and 60 to 80 commands to 4 TV cameras and 2 TV monitors. The best test runs have shown 96 to 100% voice recognition accuracy. The main conclusions of the tests are: (i) the application concept offers potential for enhancement of Shuttle operations; (ii) additional development is needed to achieve operational accuracy and reliability over a broad user population; (iii) the use of computer-recognized voice commands can contribute to a better man-machine system interaction; (iv) human acoustic characteristics and training have a major impact on system performance. As a conclusion it was decided to conduct further application tests and to promote the development of a prototype flight voice command system for future Space Shuttle applications.

I. INTRODUCTION

Efficient on-line decision making for manipulator control requires that the operator have an easy access to the relevant information sources. This is particularly important when the task requires frequent changes in the setting of a video system which contains several TV cameras and monitors in order to obtain the necessary information for manipulator control. In a fully manual control mode, where both the manipulator and video system are manually controlled, the operator can often attend either the video system control panel or the manipulator hand controllers. He cannot do both at one time. This is equivalent to a strictly sequential hand control of the manipulator and video system.

Altogether seven TV camera mounting locations exist in the cargo bay and on the manipulator of the Space Shuttle. Two TV monitors, located in the Shuttle cockpit, can be used in a split screen mode. Hence, up to four scenes can be displayed at one time. From the TV control panel in the Shuttle cockpit any camera can be linked with any monitor, and the pan, tilt, focus, iris, zoom and some internal electronic parameters of the cameras can be controlled. The control panel contains altogether thirty three pushbuttons and switches. (Figure 1.)

The RMS (Remote Manipulator System) operator normally uses both hands to control the motion of the Shuttle manipulator as shown in Fig. 2. The left hand controls the three translational motions, the right hand controls the three orientation motions of the manipulator. The video system control keyboard is under the left arm of the operator. (A few keyboard switches and pushbuttons are visible in Fig. 2.)

When simultaneous manual operation of the RMS and video system is impractical, the manual control of the Shuttle video system requires the execution of a complex multi-step process:

a. Decide which TV camera and monitor should be changed and how.
b. Stop manipulator motion, set RMS brakes on, and take hands off the manipulator hand controllers.
c. Turn visual attention to the video system control keyboard.
d. Find the appropriate buttons and switches on the keyboard.
e. Activate the appropriate buttons and switches and verify the success of this action on the keyboard.
f. Turn visual attention back to the TV monitors.
g. Verify the success of the desired information change on the monitors; if not satisfied repeat the process from step c. If everything is all right, proceed with step h.
h. Release the brakes, put hands back to the manipulator hand controllers, and continue the control task.

This process causes a disruption of RMS motion, diverts the operator's visual attention and manual work, and distracts his mental concentration from the manipulator control tasks. All these can contribute to lengthening the whole operation and to increasing operator workload.

The complex process of manual control of the Shuttle video system during manipulator operations can be considerably simplified by using a computer-based discrete word voice command system for controlling the TV cameras and monitors. Since, in effect the buttons are "pushed by voice" and the switches are "turned on/off by voice", the entire video system control process is reduced to the following simple steps:

a. Decide which TV camera and monitor should be changed and how.
b. Say the appropriate word(s).
c. Verify the success of the desired information change on the monitors, and proceed with the manipulator control task if everything is all right, otherwise repeat step b.

It can be hypothesized that voice control of the TV cameras and monitors does not disturb the operator's visual attention and manual control work, and minimizes mental distraction from the control task. Consequently, the potential of voice control for enhancing the Shuttle RMS operation was investigated in this experimental study.

A pilot voice control system was developed at JPL to test and evaluate the feasibility and utility of controlling the Space Shuttle video system by computer-recognized voice commands during manual control of the Shuttle manipulator. The voice control system is briefly described in Section II. Alternative control vocabularies are presented in Section III. Control tests conducted at the JSC MDF using the simulated full-scale Space Shuttle manipulator are described in Section IV. The test results and conclusions are summarized in Section V.

II. VOICE CONTROL SYSTEM DESCRIPTION

The pilot voice control system developed at JPL to demonstrate and evaluate Space Shuttle application concepts utilizes VDETS, a commercially available discrete word speech recognizer. VDETS is essentially a trainable acoustic pattern classifier that produces a digital code as an output in response to an input utterance. VDETS is implemented in a Nova 2 minicomputer.

The basic software used in conjunction with VDETS includes a LINC Tape Operating System (LTOS) and the VOICE Executive. LTOS allows one to edit programs and save them on a LINC tape, to store voice reference templets on a LINC tape, and to execute Nova machine language programs. The VOICE Executive is a Nova machine language core-image program that assembles user VOICE programs into Nova machine code with embedded calls to the VOICE Executive. The VOICE programming language allows one to define and develop application vocabularies and syntaxes and to perform training and recognition. The VOICE Executive is completely interrupt driven to accommodate real time response to external events.

The voice control system must be trained to each individual operator whose voice pattern templets are then stored on LINC tape for recall before using the system in the recognition mode. Training typically consists of repeating the vocabulary words set seven times as it is displayed on the self-scan display unit. The operator wears a headset with a noise cancelling microphone and adjusts the volume control to accommodate his normal speaking voice. In the recognition mode, the self-scan display shows the word recognized by the system in response to the operator's utterance.

The voice command system was connected to the TV camera and monitor control circuits through a programmable interface for which a Motorola 6802 microprocessor was employed. Whenever an operator said a command word, the programmed VDETS would send an ASCII code to the interface. The interface microprocessor would then send the data out over a parallel line to a hardware decoder which energized one of the 52 wires connected to the video system control circuits. The 6802 microprocessor also performed some simple timing and logic functions. For example, some of the switches are momentary contact switches, while the camera movement toggle switches must be held in the "on" state until a "stop" command is heard.

The video system voice control was implemented so that the commands voiced by the operator did not require verification before execution; they were executed immediately. The effect of a misrecognized command was immediately visible on the monitor. The operator needed only to voice new commands to correct for misrecognition.

The voice control system ran in parallel with the manual control keyboard so that, if required, the operator could always revert to the manual control of the video system. The main elements of the voice control system together with the overall system implementation are shown in Figs. 3-4. Performance was recorded on a printer.

III. ALTERNATIVE CONTROL VOCABULARIES

Several different combinations of vocabulary words both with and without syntax restrictions were developed and tested. Figure 5 shows a vocabulary and syntax which closely follow the words and organization of the keyboard. As seen in Fig. 5, the actual TV camera and monitor control words are arranged in five groups corresponding to the grouping of buttons and switches of the keyboard shown in Fig. 1.

In general, the syntactic organization of command words serves the purpose of increasing word recognition accuracy. The syntactic organization limits the number of words to a subset of the total vocabulary that the speech recognition system has to look up for identification of a spoken command word. Figure 6 shows a vocabulary with a multilevel syntax. As seen in Fig. 6, one can construct many subsets of the vocabulary which only contain two, three or five words. But increased syntactic grouping of words increases the application rules that the operator must remember and follow. Note also in Fig. 6 that some of the subset words are very short, e.g., "far", "in", "out", etc. Very short words have higher misrecognition probability than the longer words. The words in Fig. 6 are "natural" in the sense that they closely follow the names or functions of the keyboard buttons and switches.

The training experiments have shown that the operators prefer simple vocabularies with minimum or no syntactic restrictions. Following this desire, two vocabularies were constructed shown in Fig. 7 and 8. Note that many vocabulary words shown in Fig. 7 and 8 are concatenated words, e.g., "zoom-in", "tilt-up", "focus-far", etc. The use of concatenated words increased recognition accuracy by 6 to 8% and provided smoother and faster operation performance. The use of a concatenated word requires only one voice command (e.g., "zoom-in") for an action instead of two words (e.g., "zoom" and "in"). But some of the words shown in Fig. 7 and 8 are rather lengthy. In some cases it was necessary for the operators to speak at an unnaturally fast speech rate to get the entire utterance within the 1.5 second window that the speech recognition system allows for each spoken word. If the utterance lasts longer than 1.5 seconds, the recognition accuracy can be poor.

The vocabulary which was used during the tests at the JSC MDF is the simplest one without syntax shown in Fig. 8. It only contains two words ("stop" or "reverse") which logically must follow the action commands like "iris-open", "pan-right", etc.

IV. CONTROL TESTS

Control tests were conducted at the JSC MDF in January 1981 to evaluate the feasibility and utility of controlling the Shuttle TV cameras and monitors by voice during manual control of the Shuttle manipulator. The task configuration chosen for the tests was that of handling a Plasma Diagnostic Package (PDP) payload mock-up by the manipulator in the Shuttle bay. The PDP was berthed to and deployed from a retention mechanism.

The task started with the manipulator holding the PDP payload mock-up above the aft cargo bay area (Fig. 9). It was then docked to the retention mechanism in the aft bay, the time recorded, then deployed from there, moved and docked to a similar retention mechanism in the forward cargo bay area. The task ended when the payload was removed back to a starting position above the cargo bay. The windows of the cockpit were blocked so that the operators were forced to rely upon the TV cameras and monitors for visual feedback from the task area. The manipulation task typically required 15 to 20 minutes.

Altogether 48 test runs were performed by four operators, 32 runs with voice control of the video system. Table 1 summarizes the average number of video system control commands in both manual and voice control modes. The average number of voice commands in Table 1 does not include the misrecognized command words. Table 1 shows that the average number of manual and voice commands varies from operator to operator. It is interesting to note that the average command number variation between operators in the voice mode is less than in the manual mode.

The control tests were performed after six, seven, eight, and nine training passes for each operator. Where all the training was done at approximately the same time, six training passes seemed to give the best results. Any more than this seemed to corrupt the training patterns. The training was performed by repeating the whole vocabulary sequentially rather than repeating each word individually. When the tests were performed on a subsequent day from the training, two extra update training passes seemed to give the best results. The standard procedure was to save seven primary training passes on the LINC tape for each operator, and then update these seven passes just before the system was used in the recognition mode during the control tests, disregarding the prior updates.

During the tests the typical mode of operation was to first position the cameras and then concentrate on payload docking. This was true even with the voice system, although near the end of the tests three operators were able to combine a certain amount of camera movement with payload movement as they became more comfortable with the system.

The voice command system was used not only to select the various cameras and monitors, but also to control the camera movement and lens parameters (pan, tilt, focus, iris, zoom). The most troublesome part of the test was to control camera movement. Here the accuracy was most important since timing is critical in order to stop the movement at the right time to achieve the desired results. In most cases the operators preferred to control camera movement in "low-rate" setting. This was also the preferred setting in manual control mode. "High-rate" setting was typically used for coarse movement control.

V. RESULTS AND CONCLUSIONS

The best individual test runs have shown a recognition accuracy from 96% to 100%. As seen in Table 2, there is relatively large recognition accuracy variation between the individual operators. Three of the four operators underwent familiarization training with the voice command system at JPL two months prior to the tests at JSC. Their recognition accuracy during the tests at JSP was consistently better than the recognition accuracy of the fourth operator who learned the use of the voice command system one day before the tests.

The two "accuracy" columns in Table 2 refer to two methods of computing recognition accuracy. In the first column the accuracy is computed without the rejected words. In the second column the accuracy is computed by taking account of the rejected words. That is, rejected words were counted as errors. Each percent number belonging to an operator in the columns of Table 2 is the result from four individual test runs.

Table 2 indicates that voice recognition accuracy also depends on the vocabulary to some extent. The vocabularies JSCN04 and JSC002 in Table 2 correspond to the vocabularies shown in Figs. 7 and 8, respectively. But, as seen in Table 2, the recognition accuracy of the best scoring operator (operator B) was insensitive to both vocabulary variation and accuracy computing method.

Several off-line recognition tests (without manipulator control) were also performed at JSC with four naive users who had never previously used a voice recognition system. Their average recognition accuracy was about 90%. It is interesting to note that among the four primary operators and four naive users there were altogether three female and five male subjects, and the average recognition accuracy of the female subjects was 8-9% higher than the average recognition accuracy of the male subjects. It is also noted that only one female was used for the on-line tests, and her recognition scores were nearly perfect, ranging from 96% to 100%. Of course, these data don't have statistical significance since the test subject population was too small.

The duration of each test run with the voice command system steadily decreased as each operator became more familiar with the system. The average time per task in voice control modes was still about 10% longer than in manual control mode during the tests which should be regarded as introductory. It is felt that this time duration average will be reversed where (i) the operators gain more experience with the voice command system and (ii) the accuracy of the voice recognition system is improved. It should be kept in mind that all operators had several years extensive experience with the manual operation of the video system. As seen in Table 3, however, there was a large variation between the average time performance of the four operators even during the manual operation of the video system.

After becoming more familiar with the system, the operators were impressed with its potential and enthusiastic about it even though they felt that the recognition accuracy should be improved. In general, there was an agreement among the operators that at least 95% average total recognition accuracy is needed with a 50-word vocabulary in order for the operators to feel comfortable with the voice command system during real-time operation. In the total

recognition accuracy the rejected words are counted as errors; see last column in Table 2.

A few interesting general remarks emerged after the tests:

1) Command words should be added to the vocabular that will (i) restore camera and monitor to the condition prior to a recognition error, and (ii) allow an operator to name a selected camera position once it has been set up so that it may be re-invoked with a single word instead of repeating a complete command sequence.

2) Though the commands voiced by an operator did not require verification before execution, the operators often felt it reassuring to look at the self-scan display of the recognized words. This display, however, should be a small device and placed very close to the TV monitors.

3) The operators would like to be able to issue commands other than "stop" or "reverse" while a camera is moving. This capability would speed up the operation.

Though the control tests were not meant to test and evaluate a particular voice recognition system, it should still be mentioned that the VDETS*) system performed very well even in the presence of acoustic and electrical noise.

The main conclusions of the test are: (i) the application concept offers potential for enhancement of Shuttle operations; (ii) additional development is needed to achieve operational accuracy and reliability over a broad user population; (iii) the use of computer-recognized voice commands can contribute to a better man-machine system interaction; (iv) human acoustic characteristics and training have a major impact on system performance. As a conclusion it was decided to conduct further application tests and to promote the development of a prototype flight voice command system for future Space Shuttle applications.

Acknowledgment

The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under NASA Contract NAS7-100. The programmable interface for connecting VDETS to the video system control circuits was developed by H. C. Primus of JPL.

Note: A seven-minute narrated movie is available which shows the control tests at the JSC MDF using voice control of the Space Shuttle video system.

*) VDETS is carried by Interstate Electronics, Anaheim, CA.

Table 1. Test Summary.

- 4 TRAINED OPERATORS
- 48 TEST RUNS (EACH 15 TO 20 MINUTES)
- 32 RUNS WITH VOICE COMMANDS

	AVERAGE NUMBER OF COMMANDS	
OPERATOR	MANUAL MODE	VOICE MODE
A	57	62
B	80	70
C	47	63
D	66	52

Table 2. Voice Command Recognition Summary Accuracy.

OPERATOR	VOCABULARY	ACCURACY W/O (%)	ACCURACY W (%)
A	JSCN04 (Fig. 7)	96	90
B		97	95
C		89	86
D		80	78
AVERAGES		91	87
A	JSC002 (Fig. 8)	92	89
B		97	97
C		86	83
D		79	72
AVERAGES		89	85

Table 3. Average Task Durations.

OPERATOR	MANUAL	VOICE
A	14:36 (Minutes:Seconds)	21:12
B	20:56	20:19
C	11:25	14:74
D	25:12	25:30

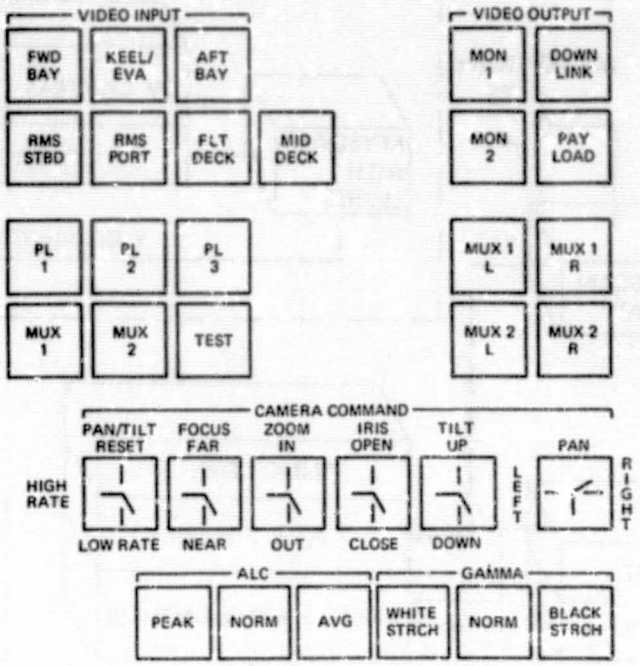


Figure 1. Space Shuttle TV Camera and Monitor Control Keyboard.

Figure 2. Space Shuttle Cockpit Control and Information Environment for Manipulator Operation.

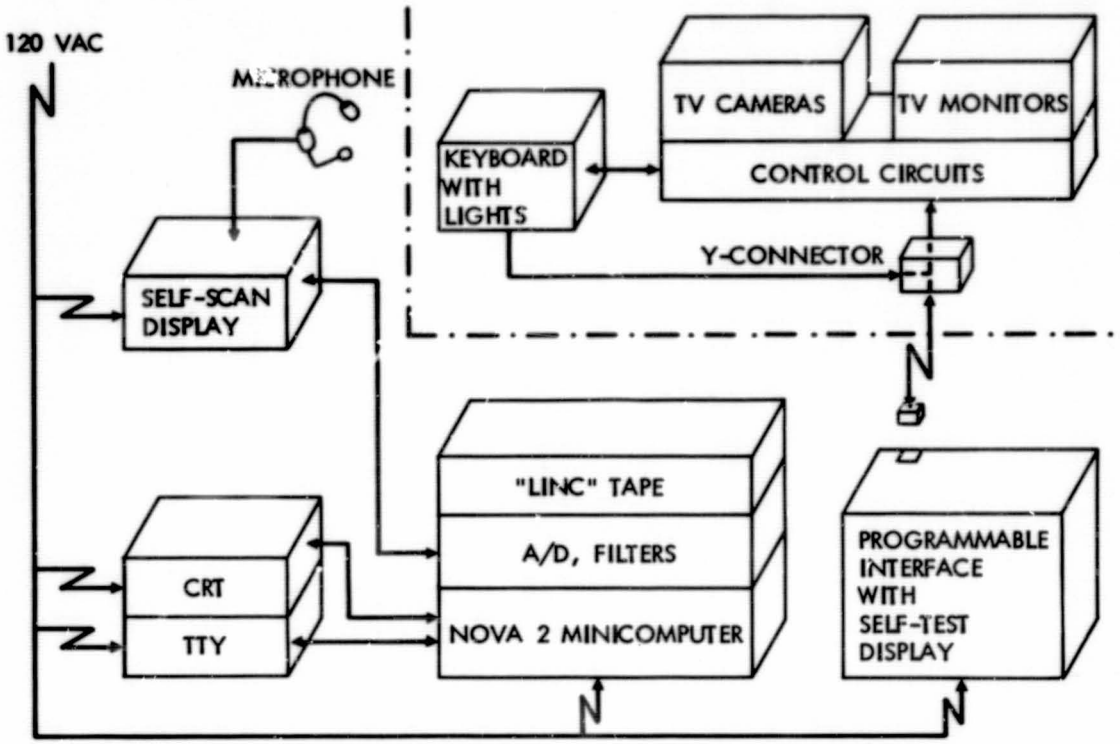


Figure 3. Main Elements of the Voice Control System and Overall System Implementation.

Figure 4. Operator Uses Voice Control of Video System During Manual Control of Shuttle Manipulator.

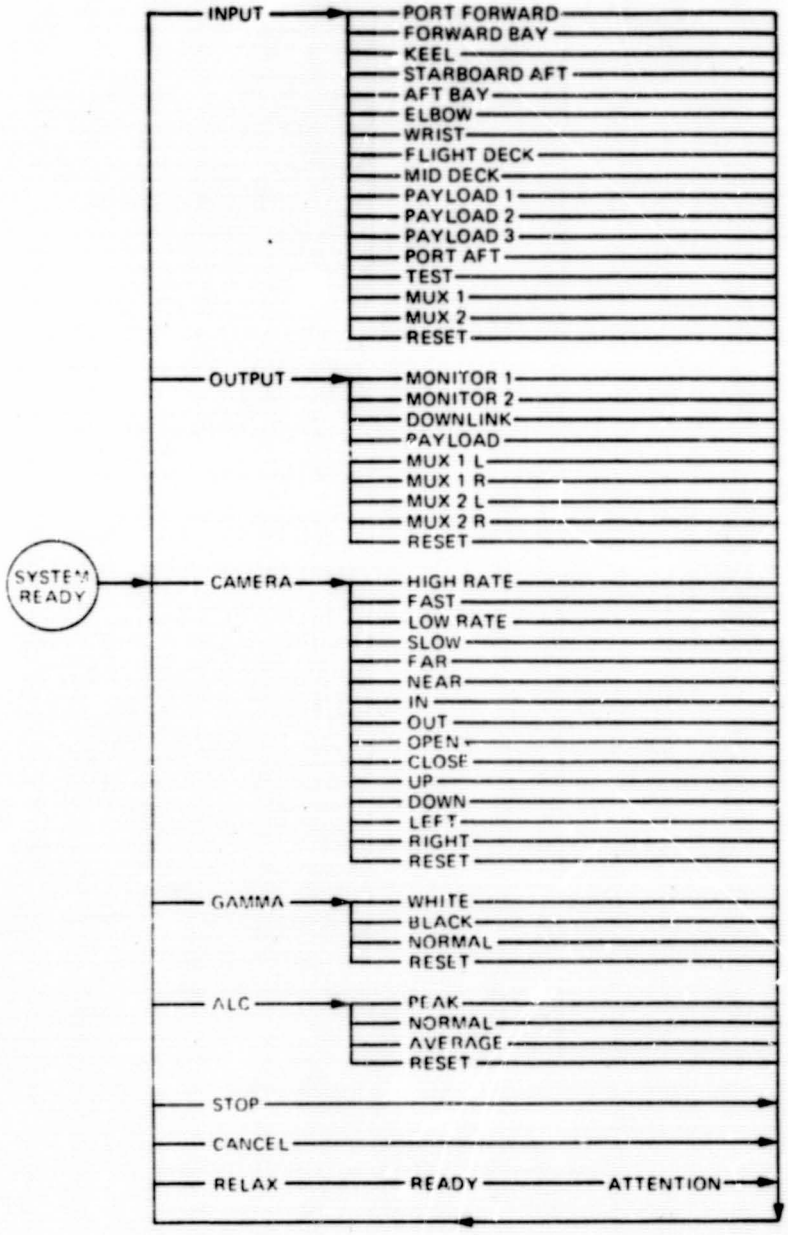


Figure 5. Natural Vocabulary with Simple Keyboard Syntax.

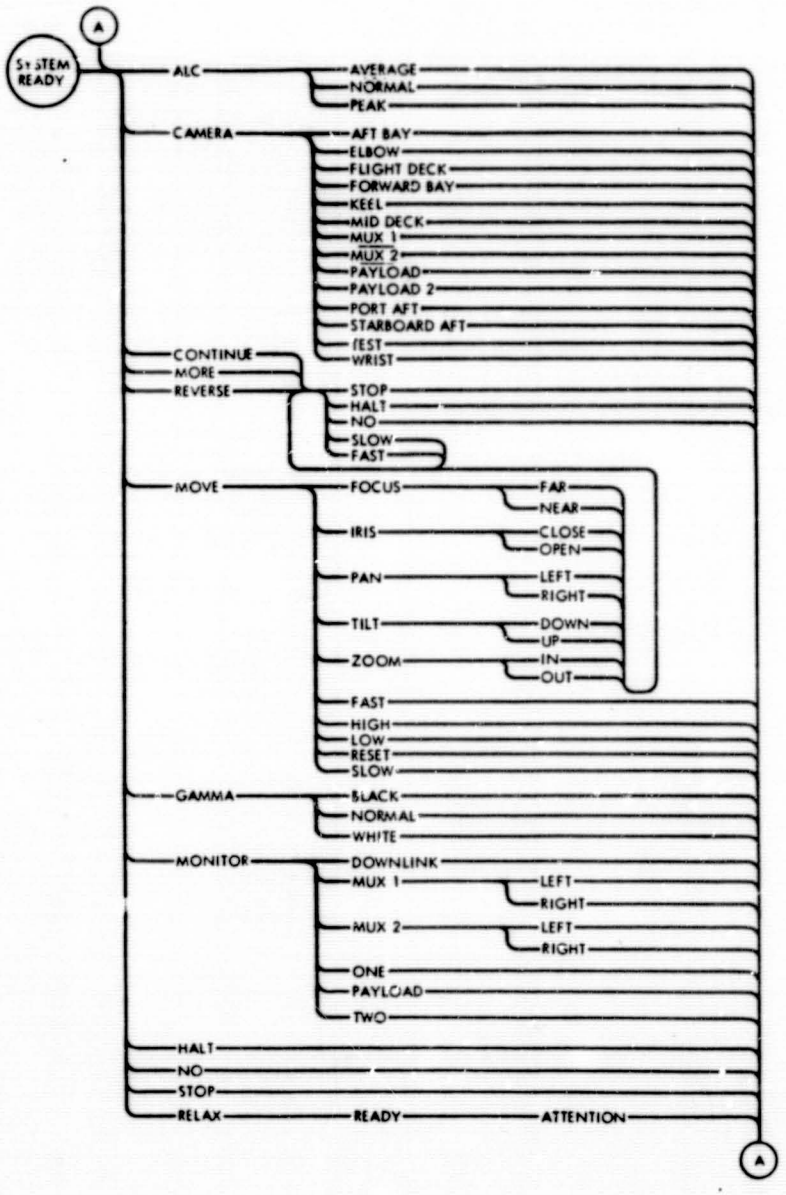


Figure 6. Natural Vocabulary with Multilevel Keyboard Syntax.

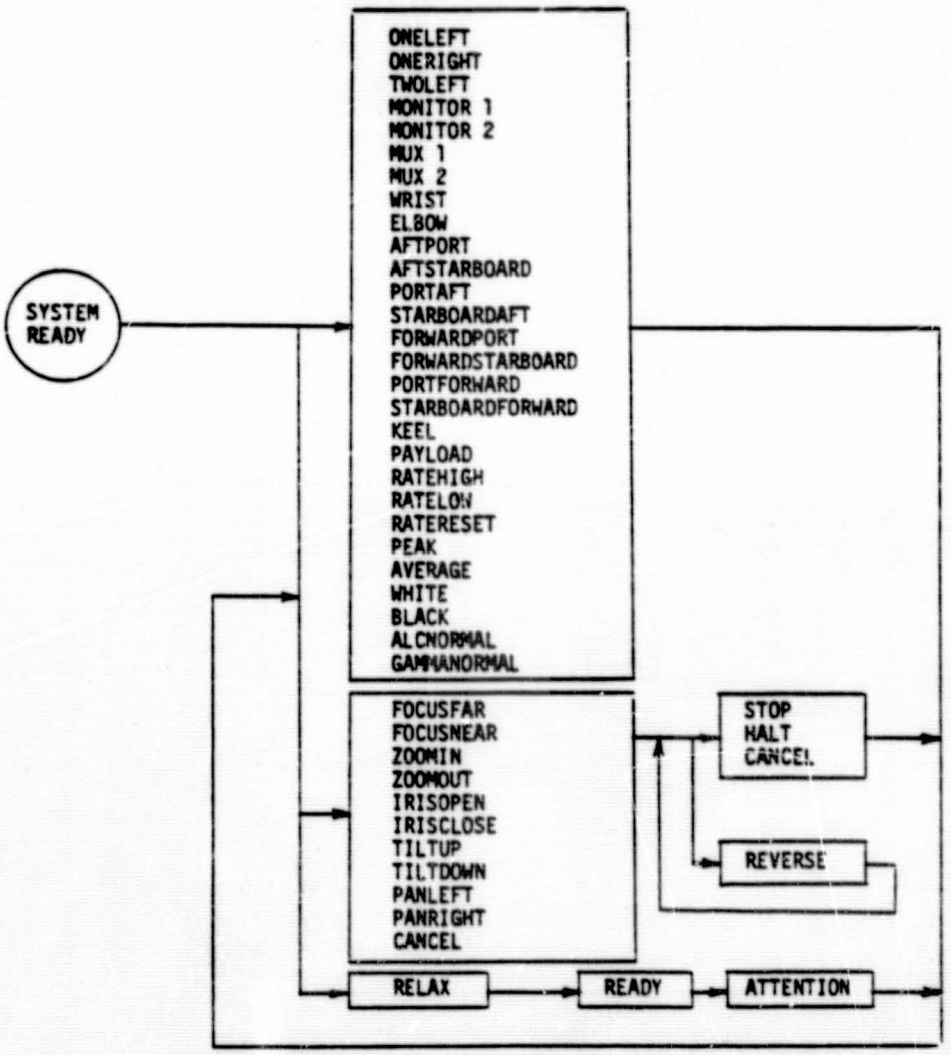


Figure 7. Vocabulary with Concatenated Words and Minimum Syntax.

PORT-FORWARD	AFT-PORT	MONITOR-1
FORWARD-PORT	PORT-AFT	MONITOR-2
		MUX-1-LEFT
AFT-STARBOARD	FORWARD-STARBOARD	MUX-1-RIGHT
STARBOARD-AFT	STARBOARD-FORWARD	MUX-2-LEFT
MUX-1	RMS-PORT	MUX-2-RIGHT
MUX-2	ELBOW	

RATE-HIGH	FOCUS-FAR	ZOOM-IN	IRIS-OPEN	TILT-UP	PAN-LEFT
RATE-LOW	FOCUS-NEAR	ZOOM-OUT	IRIS-CLOSE	TILT-DOWN	PAN-RIGHT

REVERSE STOP

PEAK ALC-NORMAL AVERAGE WHITE GAMMA-NORMAL BLACK

RELAX READY ATTENTION

Figure 8. Reduced Vocabulary with Concatenated Words and without Syntax.

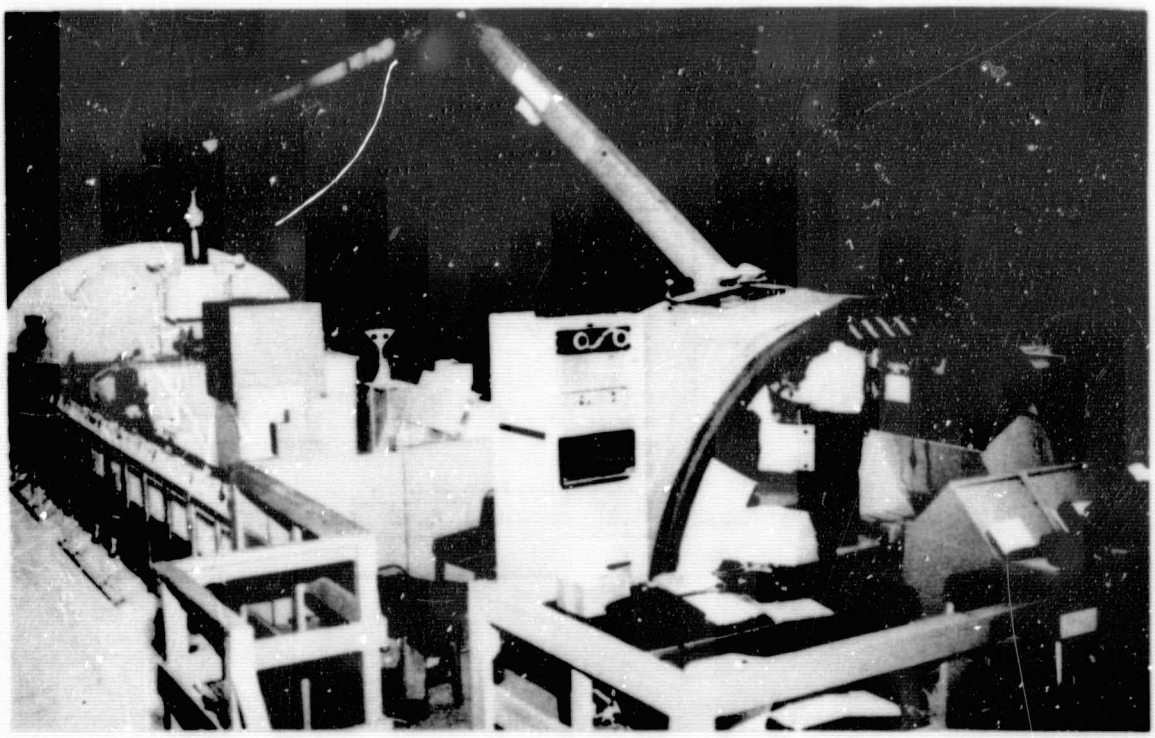


Figure 9. Task Scene for Voice Control of the Shuttle Video System During Manual Control of the Shuttle Manipulator.

A Comparison of Head and Manual Control for a Position-Control Pursuit Tracking Task

by

William H. Levison and Greg L. Zacharias
Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, MA 02238

and

James L. Porterfield, Donald Monk, and Christopher Arbak
Air Force Aerospace Medical Research Laboratory
Wright-Patterson AFB
Ohio 45433

ABSTRACT

Head control was compared with manual control in a pursuit tracking task involving proportional controlled-element dynamics. An integrated control/display system was used to explore tracking effectiveness in horizontal and vertical axes tracked singly and concurrently. Compared with manual tracking, head tracking resulted in a 50% greater rms error score, lower pilot gain, greater high-frequency phase lag and greater low-frequency remnant. These differences were statistically significant, but differences between horizontal- and vertical-axis tracking and between 1- and 2-axis tracking were generally small and not highly significant. Manual tracking results were matched with the optimal control model using pilot-related parameters typical of those found in previous manual control studies. Head tracking performance was predicted with good accuracy using the manual tracking model plus a model for head/neck response dynamics obtained from the literature.

Objectives

An extensive research program is underway in the Air Force to develop predictive models for pilot control behavior for use in the design of advanced aircraft and ground-based flight simulators. Such models must be applicable to a variety of task environments, including (a) steady-state and non-steady-state control problems, (b) cockpit instruments and extra-cockpit visual scenes, and (c) manual and head control modes. Use of the head as a control effector is of particular relevance to tasks, such as airborne weapons delivery, where the operator may be required to perform multiple control tasks.

This paper summarizes the analytical work performed by Bolt Beranek and Newman Inc. (BBN) in support of an experimental study conducted by the Air Force to compare head and manual control performance. Additional

details are documented in [1]. Model analysis was performed with the optimal control model (OCM) of the human operator. This model was selected because of its demonstrated predictive capabilities across a broad spectrum of task environments [2]. As shown later, the model provides a consistent treatment of the head and manual control results obtained in the Air Force study.

Background

The use of the head as a control effector in continuous tracking tasks has not been explored to a great extent. Studies of head tracking have not, in general, used the full range of performance assessment techniques often employed in studies of manual control, nor have they proposed or validated mathematical models. Furthermore, experimental results run counter to what we would expect on the basis of manual control results.

A comparison of head and manual tracking is provided by Chouet and Young in a study employing both time- and frequency-response performance measures [3]. A set of spatial orientation tasks were performed which required the subjects to regulate the attitude of a moving simulator cab in the presence of a pseudo-random disturbance input. Rate control of the cab was implemented.

Head tracking compared favorably on the average with manual control in the pitch and yaw axes but was less effective in the roll axis. Compared with manual tracking, the gain crossover frequency* for head control was the same in the yaw axis and about 20% less in the pitch axis. Integral squared tracking error, averaged over these two axes, was about 16% greater for head tracking.

Shirachi, Monk, and Black [4] studied the head control effector in a proportional-control pursuit tracking task. Control was performed singly and jointly in the horizontal and vertical axes. Manual control was not explored, and only frequency-response measures were obtained.

Differences between axes and between 1- and 2-axis conditions were found. Pilot gain was substantially greater in the vertical axis, whereas pilot response was more highly correlated with the tracking input in the horizontal axis. Dual-axis tracking appeared to be more efficient than single-axis tracking: specifically, pilot gain and response coherency were greater, and phase lag was smaller, for the dual-axis task.

A number of the findings reported in these two studies are surprising in light of other studies of human controller behavior. First, since the

---

* The gain-crossover frequency is the frequency at which the combined transfer characteristic of the operator and controlled element is unity (0 dB).

hand is less massive than the head, one would expect manual control to be of wider bandwidth -- and thus more effective -- than head control. Second, previous manual control studies have shown similar tracking effectiveness in vertical and horizontal axes when the control tasks have been statistically identical on the two axes [5,6]. Finally, studies of multi-axis manual control have shown that performance on a given axis either stays about the same or degrades when another axis is tracked concurrently; it does not tend to improve. The experimental program summarized in this paper was conducted, in part, to resolve these discrepancies as well as to develop and test a predictive model for head tracking.

As an initial working hypothesis, we adopted a model of head/neck dynamics similar to that proposed by Merasso et al [7], which was based on the response of the head to passive displacement. They fitted the observed response with a second-order system having a natural frequency of 9 rad/sec and a damping ratio of 0.55. They also added a 20 msec delay and a first-order low-pass filter having a time constant of 0.18 sec to reflect additional neuromuscular response mechanisms.

DESCRIPTION OF EXPERIMENT

An experimental program was conducted to explore the ability of the human operator to perform a pursuit tracking task using the head as the control effector; manual tracking with a nearly isometric control stick was also explored to provide a point of reference. The output of the controlled element ("plant") was proportional to the operator's control input. Plant position and target displacement were explicitly displayed to the operator. Tracking was performed in three modes: (1) horizontal axis only, with vertical error clamped electronically at zero; (2) vertical axis only, with horizontal error clamped at zero; and (3) combined horizontal and vertical axes.

Half the experimental trials were performed with the subject controlling the cursor by appropriate head movements. A helmet-mounted sight was used to sense the subject's head angles as he tracked the target and was calibrated electrically so that one degree of head rotation produced one degree of cursor displacement. A nearly isometric control stick was used as the manual input in the remaining trials. In order to facilitate comparison of pilot response characteristics obtained in the two control modes, system gains were adjusted so that one volt of recorded control input always corresponded to 1 degree of cursor displacement. One control volt represented one degree of head motion, or 1/8 lb control force.

Forcing functions were constructed from 11 sinewaves whose amplitudes were selected to simulate a white noise process passed through a second-order filter having a double pole at 2 rad/sec. Frequencies were

spaced at approximately half-octave intervals. Horizontal-axis input frequencies were interleaved with vertical-axis frequencies to allow for differentiation between horizontal-axis and vertical-axis response characteristics during 2-axis tracking.

Eight university students served as subjects for this experiment. Each subject served in all six conditions of control mode (head or manual) and target motion (horizontal only, vertical only, and 2-axis). Half the subjects were first trained and tested with the joystick and were then trained and tested on the head motion system; half were trained in the reverse manner.

A tracking session consisted of three sets of four 100-sec trials, one set with each type of target motion. The first 9 seconds of each run were considered as "start up" time; the remaining 91 seconds were recorded and scored. A 1-minute rest period was provided between each trial within a set of four trails, and a 5-minute rest was provided between each set. Subjects were instructed to minimize the circular error probability (CEP)* and were given their CEP scores at the end of each run. Order of presentation was counterbalanced over subjects.

Each subject was trained until a performance asymptote was reached, where "asymptote" was defined as an improvement of 5% or less averaged over all trials in a session on two consecutive days. On the average, the subjects received about 100 practice trials total. Experimental data were taken on the day following the day a subject reached asymptote. Each subject provided 24 trials of experimental data: 2 control modes, 3 target conditions, 4 replications each.

EXPERIMENTAL RESULTS

The analysis procedure followed in numerous previous studies was again followed here. First, the time histories were analyzed to provide various time- and frequency-domain measures of tracking performance. Second, these results were averaged across subjects and then subjected to model analysis to identify (i.e., quantify) parameters relating to operator response limitations. Emphasis was placed on testing a predictive model for head tracking.

Primary Data Reduction

Tracking error and control input time histories from each experimental trial were subjected to fast-Fourier transform (FFT) analysis to yield,

---

* The CEP score was defined as the radius of an imaginary circle drawn around the target such that the cursor was within this circle 50% of the time. For a Gaussian tracking error distribution, the CEP is proportional to rms tracking error.

among other measures, estimates of power spectra. These spectra were then integrated to obtain estimates of error and control variance. These same spectra were also multiplied at each FFT frequency by the square of the frequency (in rad/sec) and again integrated to obtain estimates of error-rate and control-rate variance.

RMS performance scores, obtained by taking the square root of the population mean for each variance score, are presented in Figure 1. Tracking error scores were nearly identical for horizontal- and vertical-axis tracking and were little influenced by the number of axes tracked simultaneously. Error scores associated with head tracking, however, were about 50% greater than manual tracking scores. Results of paired t-tests performed on variance scores, reveals that head/hand differences in tracking error were statistically significant at the 0.001 level. Head/hand differences in other rms performance measures were inconsistent and generally not statistically significant.

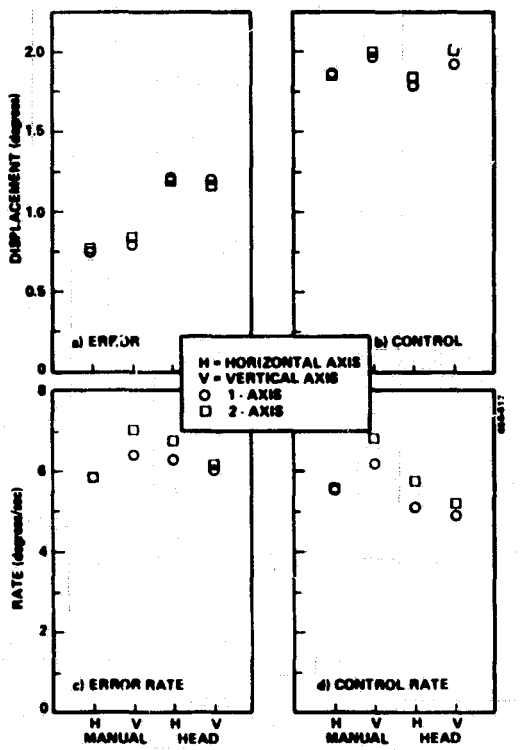


Figure 1. RMS Performance Measures
Average of 8 subjects, 4 trials/subject

Horizontal/vertical differences on the order of 5 to 20% were observed for the remaining three performance measures. For the most part, however, these differences (as well as 1-axis/2-axis differences) were not strongly significant.

Results of the FFT analysis were used to compute estimates of the pilot describing function and remnant spectrum for each experimental trial. Two sets of measures -- one for each axis -- were obtained for 2-axis trials. The describing function was expressed in terms of amplitude ratio ("gain") and relative phase shift, whereas the portion of the closed-loop control spectrum not linearly correlated with the tracking input served as the measure of pilot remnant.

Amplitude ratio and phase measures were very nearly similar for the horizontal and vertical axes and for single- and dual-axis tasks (see Levison et al [1]). The frequency dependency of the remnant spectrum was virtually unchanged, but the magnitude was about 1-2 dB greater for vertical tracking. The remnant spectrum (for a given axis) was also about 1-2 dB greater for dual- than for single-axis tracking. Thus, the small differences in tracking performance related to axis orientation and to number of concurrent axes appears to stem from differences in the "noisiness" of the pilot's response.

Considerably greater differences in tracking performance were associated with the mode of tracking. Figure 2 shows that, in comparison to manual tracking, head tracking yielded lower amplitude ratio and higher remnant at low and mid frequencies, and larger phase shift at high frequencies. These differences were highly significant [1]. As shown by the model analysis below, the differences shown in Figure 2, as well as the head/hand differences in performance scores discussed above, are interrelated and reflect a consistent cause-and-effect relationship.

Model Analysis

The Optimal Control Model (OCM) of the human operator was employed to provide a theoretical framework to account for the effects of control mode (head or hand) on pilot response behavior. Our objective was to seek a consistent modeling philosophy that would replicate manual and head tracking performance. Readers unfamiliar with this model are directed to the review article by Baron and Levison [2], and the citations listed therein, for a description of the model structure and parameterization.

Because control mode was the only experimental variable to yield performance differences that were significant in both the practical and statistical sense, model analysis was directed toward explaining head/hand differences. Average data obtained for the single-axis horizontal tracking task were used to identify pilot-related model parameters and to test the predictive capability of the model.

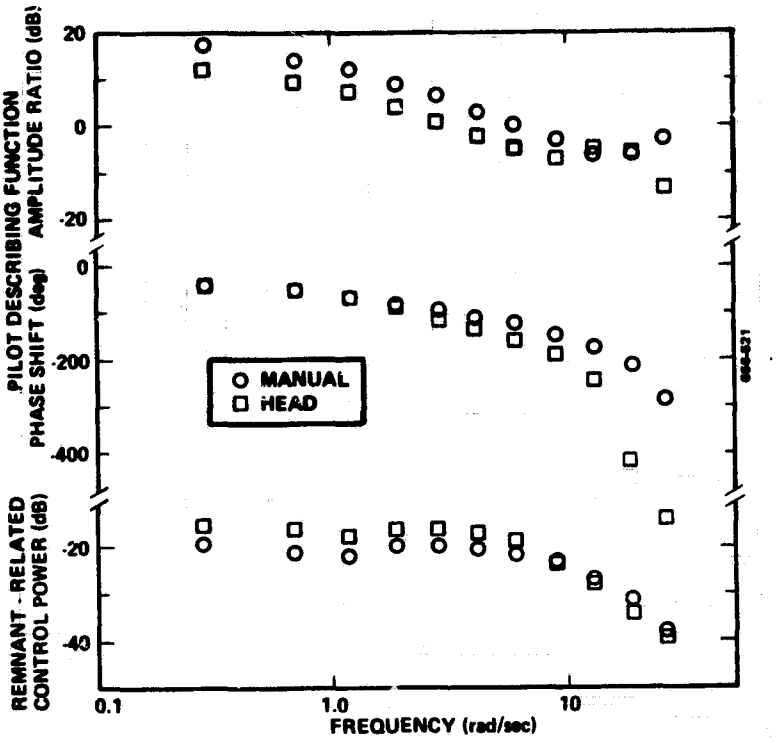


Figure 2. Effect of Tracking Mode on Frequency Response
1-axis horizontal tracking.
Average of 8 subjects, 4 trials/subject

Parameters for the manual control task were first identified using the gradient search scheme reported by Lancraft and Kleinman [8] and modified by Levison [9]. This scheme required the definition of a scalar "matching error" that consists of normalized squared differences between model predictions and experimental measurements. Variance scores, describing function gain, describing function phase shift, and pilot remnant measurements were used in computing this matching error.

Each model-data mismatch was normalized with respect to the (across-subject) standard deviation of the experimental mean so that: (a) a set of dimensionless quantities would result, allowing their accumulation into an overall scalar matching error, and (b) each component matching error would contribute to the total in proportion to the reliability of the data.

This gradient search scheme identified the parameter values shown in the first column of Table 1. These values were then "rounded off" as shown in the second column of Table 1 and tested against the same matching criterion. As the resulting matching error was within 12% of that obtained by the gradient search, the latter set of parameter values were used to model the head tracking data.

A model for the head/neck system similar to that proposed by Morasso et al [7] was tested against the head tracking data obtained in this experimental study (1-axis, horizontal task). To model this task, the system dynamics were augmented by a second-order filter having a natural frequency of 9 rad/sec and a damping ratio of 0.55, plus a first-order lag of 0.18 seconds. The output of this third-order system was considered as the operator's control signal for purposes of predicting the pilot describing function and remnant spectrum.

This modified pilot model was tested against the experimental head tracking data with pilot parameters adjusted as indicated in the second column of Table 1. As shown in Figure 3, a good match between model and experiment was obtained over much of the measurement bandwidth.

Experimental and predicted rms performance scores are compared for both the manual and the head tracking tasks in Figure 3a. Brackets indicate one standard deviation as measured across subjects. The "approximate" pilot parameters shown in Table 1 were used for both comparisons. Because the model results for the head tracking tasks are based on parameters identified for the manual task, plus a model for head dynamics taken from the literature, these results are true predictions.

All four rms performance scores obtained from the manual control data were matched to within one standard deviation. Although the head-tracking scores were predicted less accurately, all predictions were within 12% of the experimental mean, and rms tracking error was predicted to within 5%.

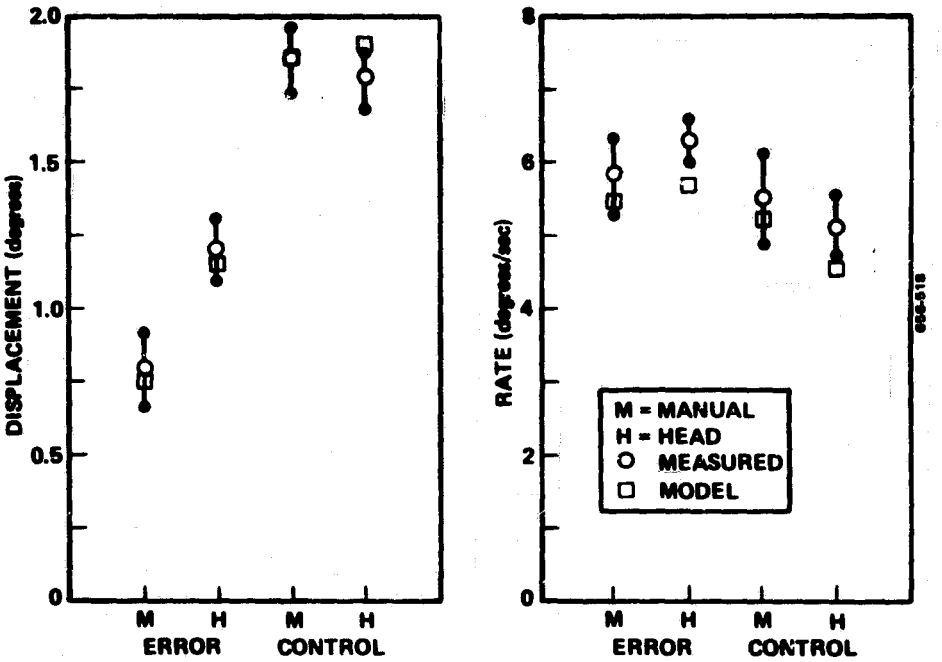
Figure 3b shows a very close match to frequency response measurements obtained from the manual tracking data. Measurements at mid frequencies -- where response behavior is critical in terms of overall system behavior -- were predicted with high accuracy for the head control task. Prediction errors were greatest for low-frequency amplitude ratio and phase shift, and high-frequency amplitude ratio and remnant.

DISCUSSION

The parameters identified for the manual task are consistent with values found in earlier studies using proportional control systems [2].*

* This is true for all parameters except motor noise/signal ratio, which cannot be meaningfully compared with previous results because of the different treatment of motor noise [1].

a) RMS Performance Scores



b) Frequency Response

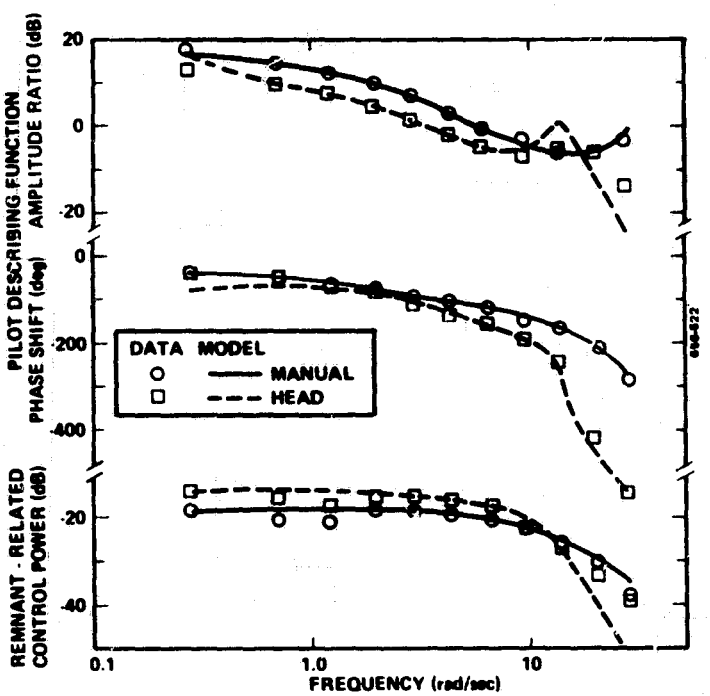


Figure 3. Comparison of Model and Experimental Results

1-axis horizontal tracking.
Average of 8 subjects, 4 trials/subject

Table 2

Identified Values for Pilot-Related Model Parameters

Parameter	Best Fit	Approximate Value
Matching error	0.840	.937
P_e	-21.0	-20
$P_{\dot{e}}$	-19.5	-20
P_u	-26.0	-25
TD	0.169	0.17
TN	0.082	0.08

P_e = observation noise/signal ratio, error (dB)

$P_{\dot{e}}$ = observation noise/signal ratio, error rate (dB)

P_u = motor noise/signal ratio (dB)

TD = time delay (seconds)

TN = motor time constant (seconds)

Single-axis, horizontal, manual control task.
Average of 8 subjects, 4 trials/subject

The ability to predict head tracking data with these parameter values, plus a model for head/neck response dynamics obtained from the literature, suggests that the optimal control model provides a mechanism for generalizing the results of this study to other tasks. Specifically, we would expect this model to be valid for tasks in which overall system performance is relatively insensitive to pilot response behavior at frequencies less than 0.5 rad/sec or greater than 10 rad/sec. Although one could probably improve the match at the high and low end of the measurement band, through readjustment of the pilot-related parameters, we submit that a more meaningful approach -- one having greater predictive potential -- would be to revise the model for head/neck response dynamics.

The reader is cautioned against generalizing head/hand differences on the basis of the performance differences obtained in this study. In general, head/hand differences can be expected to depend on the details of the task environment, including controlled element dynamics, external

forcing function characteristics, and possibly performance requirements. It is these potential interactions that provide the primary motivation for model development.

The lack of 1-axis, 2-axis differences reported here are consistent with earlier studies involving "integrated" controls and displays which employ (a) a single manipulator having similar characteristics in two dimensions, and (b) a single error indicator that moves in two dimensions [5]. Were the display to be non-integrated such that separate display elements indicated horizontal and vertical tracking error, significant 1-axis, 2-axis performance differences would be expected [6].

We cannot explain the differences between the results of this study and some of the counterintuitive results reported in previous studies of head tracking; published information is inadequate to allow complete reconstruction of the earlier studies. We can only point out that care was taken in this study to provide the subjects with knowledge of performance after each practice trial, and to train them until apparent asymptotic performance. We assume that this training procedure motivated the subjects to minimize their error scores whatever the task conditions.

REFERENCES

1. Levison, W.H., and G.L. Zacharias, "A Comparison of Head and Manual Control for a Position-Control Pursuit Tracking Task", Report No. 4572, Bolt Beranek and Newman, Inc., Cambridge, MA, January 1981.

2. Baron S. and W.H. Levison, "The Optimal Control Model: Status and Future Directions", Proceedings of the 1980 International Conference on Cybernetics and Society, Cambridge, MA, Oct. 8-10, 1980.

3. Chouet, B.A. and L.R. Young, "Tracking with Head Position Using an Electrooptical Monitor", IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-4, No. 2, March 1974, pp. 192-204.

4. Shirachi, D.K., D.L. Monk and J.H. Black, Jr., "Head Rotational Spectral Characteristics During Two-Dimensional Smooth Pursuit Tasks", IEEE Transactions on Systems, Vol. SMC-8, No. 9, Sept. 1978, pp. 715-724.

5. Levison, W.H., "Two-Dimensional Manual Control Systems", Proceedings of the Second Annual Conference on Manual Control, Cambridge, MA, Feb. 1966, pp. 159-181.

6. Levison, W.H., J.I. Elkind and J.L. Ward, "Studies of Multivariable Manual Control Systems: A Model for Task Interference", NASA CR-1746, May 1971.

7. Morasso, P., G. Sandini, V. Tagliasco, and R. Zaccaria, "Control Strategies in the Eye-Head Coordination System", IEEE Transactions on

Systems, Man, and Cybernetics, Vol. SMC-7, No. 9, Sept. 1977, pp. 639-651.

8. Lancraft, R.E. and D.L. Kleinman, "On the Identification of Parameters in the Optimal Control Model", Proceedings of the Fifteenth Annual Conference on Manual Control, Ohio, Nov. 1979, pp.487-502.

9. Levison, W.H., "A Comparison of Head and Manual Control for a Position-Control Pursuit Tracking Task", Proceedings of the 17th ANnual Conference on Manual Control, June 1, 1981.

GMT

FIRE CONTROL SYSTEMS EVALUATION — THE MAN-MACHINE PERSPECTIVE

by

Dr. J. Korn

ALPHATECH, Inc.
3 New England Executive Park
Burlington, Massachusetts 01803
(617) 273-3388

SUMMARY

In this paper we analyze the tracking and gun-pointing performance of two generic fire control/gunner systems (FCS), the disturbed reticle (DR), and the stabilized sight-director (SS). This analysis is facilitated by using the Optimal Control Model (OCM) concept. The DR system constitutes a representative design of an FCS actually implemented on the Chrysler XM-1 tank, whereas the SS is a proposed alternative design. Both systems include a (first-order) gun-lead prediction mechanism, which is designed to compensate for the time-of-flight of the fired projectile, but they differ in the interaction among their respective components. In the DR design, the functions of target tracking, lead-prediction, and gun-pointing are intermixed in a complex dynamic interaction of the gunner, sight (visual loop), and gun-turret. On the other hand, the SS system components are dynamically separated such that the pointing mechanism does not disturb the visual loop.

The visual loop dynamics are similar in both systems (~K/s), thus effecting comparable tracking performance. We then define the gun-pointing error as the error between the target position (LOS angle) T seconds ahead, where T = projectile time-of-flight, and the current gun-pointing angle. It is shown that the DR exhibits a sluggish and inferior pointing response when compared to the SS system, and therefore, it is determined that the gunner is not the primary source of degraded pointing performance.

We further show that a first-order lead prediction is inadequate, in general, in the case of a maneuvering target. Therefore, a higher-order prediction algorithm in conjunction with the SS system is proposed. This prediction system uses range measurement and the gunner-commanded pointing angle in estimating the target state (position, velocity, and acceleration). This estimate is then used in the appropriate (optimal) second-order predictor.

Preliminary results show superior pointing performance, and suggest the applicability of a higher-order prediction mechanism in fire control systems.

INTERRUPTION AS A TEST OF THE USER-COMPUTER INTERFACE

By John G. Kreifeldt and Mary E. McCarthy*

Dept. of Engineering Design
Tufts University
Medford, MA. 02155

17th Annual Conference on Manual Control

SUMMARY

In any practical interaction with a computer, the user is required to formulate, phrase and enter the problem for solution through a series of steps. One very common and important characteristic of this procedure is that the operator may be interrupted at any point and required to attend to other tasks before resuming. Such interruption could have pronounced effects on the time and errors made while completing the computer task.

In order to study the effects different logic systems might have on interrupted operation, an Algebraic calculator and a Reverse Polish Notation calculator were compared when trained users were interrupted during problem entry. The RPN calculator showed markedly superior resistance to interruption effects compared to the AN calculator although no significant differences were found when the users were not interrupted.

Causes and possible remedies for interruption effects are speculated. It is proposed that because interruption is such a common occurrence, it be incorporated into comparative evaluation tests of different logic system and control/display system and that interruption resistance be adopted as a specific design criteria for such design.

INTRODUCTION

Designing the user-computer interface has been a decided challenge and will continue to assume growing importance. One of the vital aspects concerns interfacing the human user's cognitive abilities and characteristics with those of the computer, i.e. its logic program.

*The work reported here is based on the unpublished M. S. thesis of Mrs. Mary McCarthy.

Grace (1970.) suggested that a new relationship should be regarded as existing within a man-machine logic system, and that the logic component introduces a new set of considerations for the human factor discipline.

Man-computer cooperation in problem solving places emphasis on the design of the console devices for efficient communication of information between man and the computer. This mainly involves the computer display and the console manipulatory devices such as the switches, light pens, and plugs. The display and console manipulatory devices serve not only to transmit information between the human to the computer but also, in varing degrees to record and establish memory of this action and its consequence. (Zeigler and Sheridan, 1965.) The form and content of the information conveyed by these devices must be such as to aid the human in his thought processes, and thus make it easier for the operator to perform multitasks more efficiently.

In 1977, Durding, Becker and Gould studied the human memory skills as to how they aid or hinder data processing. They considered that memory organization depends on the task condition (computer memory) and if this were optimized then the human memory would be used more efficiently. However, at best, in the real world situation, the human mind seldom is without distraction, as for example unexpected interruptions. Enstrom and Rouse in 1977, did considerable work in analyzing how the human's allocation of attention will affect the input-output of the human relative to the computer.

It is obvious that the display and devices of the data processing machines immediately effect the transfer of information, yet there is a more fundamental area where man and machine meet, where the reasoning powers of man interact with the very essence of the machine, its logic system.

It is clear that this logic interfacing is less than ideal. However, one of the most significant challange is to specify useful design principles and evaluation criteria for logic design in addition to its functional characteristics in a manner similar to specifying physical characteristics such as button placement, size, etc., relating to proper computer hardware design for human use.

Common (and common sense) criteria for evaluating alternative logic design at the user level ordinarily include the number of key strokes required, number of errors made, and the time required to enter the problem and obtain the results.

Card (1979) found no difference between RPN (Reverse Polish Notation) and AN (Algebraic Notation) calculators on the basis of the time required to enter and solve problems of various degrees of complexity. However, a recent study (Agate & Drury, 1980) found significant differences in percentage errors and average completion times between AN and RPN with RPN showing the more favorable performance for both measures. On the other hand, a different study (Kasprzyk, Drury, Bialas, 1979) found smaller significant differences between AN and RPN with RPN still superior.

One of the testing principles in evaluating alternative designs that do not show clear performance differentiation is to apply increasing amounts of user or environmental stress, simulating conditions which could reasonably arise, and compare the alternative design performances. The design maintaining a higher level of performance over the wider range of stress could reasonably be judged the superior design. The assumption holds that the user is taxed less by the more stress-tolerant design, which would seem to be a sensible design criteria (Chapanis, 1959).

This principle of user stress was applied to a performance re-evaluation of an RPN and an AN hand calculator. The purpose of the experiment was not to evaluate particular calculators per-se but rather to devise a test methodology for evaluating different logic system designs in a practical, user-oriented manner. The calculators were used because they provided existing (and presumably optimized) examples of alternative logic designs for solving identical problems.

The "stress" applied to the user simulated real world conditions often encountered in which the problem solving task is interrupted part way through for some reason, with a subsequent diversion of attention. After the interruption, the operator is required to resume the task. Examples of this type of interruption are familiar to pilots, controllers, and others in a multi-task environment. A more homely example might be that of a housewife programming a microwave oven, being interrupted by a child or other distractions and then resuming the programming. In fact, anyone using a hand calculator has commonly experienced being interrupted during a lengthy series of key strokes with subsequent consequences.

The issue raised by the interruption is its impact on task completion. It is a reasonable assertion that a superior design would, among other virtues, show the least difference on the primary entry task between interrupted and uninterrupted performances. Common effects of an interruption include causing the user to loose the place, and/or to start over.

Thus a "good" logic design should, in principle, permit the user to resume the task after interruptions with negligible effects on subsequent performance. A corollary principle is that the user should be able to reprogram at any point making maximum use of previously programmed material. An example of this could be found among pilots programming an onboard navigational computer through a series of waypoints with associated altitudes, speeds and times. If interrupted during this task, the pilot should be able to resume as though no interruption occurred and/or, if necessary, reprogram at any time to meet new requirements. A clearly inferior logic design would require the pilot to begin again from the beginning effectively wasting any time previously spent on programming material which remains a subject of the existing problem.

There are at least two major components to a logic system for human use. One component is the logic system which on hand calculators manifests itself as the manner in which a problem must be phrased for entry (e.g. Reverse Polish Notation or Algebraic Notation). The other component is the console unit of input and display systems whose purposes include

verification, feedback, and resultant display.

One could assume that the logic system should be compatible with the user's own "logic system" or cognitive characteristics of problem phrasing, whether learned or innate, while the console system should at least aid during the process of problem entry particularly if information processing is required (Ziegler, Sheridan, 1965). Except for very novice users, problems are never entered from a prewritten list of key strokes (except possible for lengthy programs entered into a programmable calculator). Thus since the hand calculator user works directly from mind to calculator, errors in solution (barring key stroke mistrikes) as well as the time taken to enter a problem (sequence of key strokes) could be attributable to mismatches between the human and machine logic systems and inadequacy of the display system. Any distraction interrupting the user during the entry process would put stress on the logic interface and on the adequacy of the console system.

The objective of our experiment was to test the effects of interruption during the calculation task on errors and completion time for AN and RPN logic systems as examples of different logic systems designed for the same purposes. The purposes of the experiment were to examine task interruption - a common occurrence - as a potential sensitive probe of performance with different logic systems and to shed further light on potential differences between AN and RPN calculators in a simulated multi-task environment. Based on the equivocal findings in performance between the two types of calculators when the user is not interrupted, the superior design would be expected to show the lesser difference between interrupted and noninterrupted performances.

METHODOLOGY

The Hewlett Packard Model 31E and the Texas Instrument Model TI57 hand calculators were chosen as examples embodying the RPN and the AN logic, respectively. Both calculators are examples of scientific calculators equipped to solve problems of considerable formulation complexity.

Four problems of different prima facie complexities were formulated to span a range of types that might be encountered. The problems were:

(1) $$\frac{5 \times (4 + (3 \times (2 - 6)))}{\frac{6 \times \sqrt{22}}{17} + \sin 37^{o}} = (-17.7206)$$

(2)
254
762
321
854
952
1859
2
400
37
369
824
592
333
1
(7460)

(3) $3 \times (4^{(2 - \sqrt[4]{7})}) = (4.7000434)$

(4) $\dfrac{1.2 \times 10^{16} + 3.45 \times 10^{14}}{(4 \div 5^{2}) \times 7 + (3 \times 5^{\cos 60^{\circ}})} = (1.577 \times 10^{15})$

The answers in parentheses were, of course, not presented to the subjects. At this point, the reader is encouraged to imagine which of the above problems, if any, might reveal the largest performance differences between the two calculators.

In order to contrast RPN with AN, the first problem is shown in Figure 1 as it would be entered into each calculator.

$$\frac{5x(4 + (3x(2 - 6)))}{\frac{6 \times \sqrt{22}}{17} + \text{SIN } 37^{\circ}} = -17.7206$$

REVERSE POLISH NOTATION			ALGEBRAIC NOTATION		
KEY	DISPLAY	COMMENT	KEY	DISPLAY	COMMENT
2	2.		5	5.	
ENTER	2.00		x	5.	Mult. pending
6	6.		(	5.	
-	-4.00	result of Subt.	4	4.	Add pending
3	3.		+	4.	Add pending
x	-12.00	result of Mult.	(	4.	
4	4.		3	3.	
+	-8.00	result of Add	x	3.	Mult. pending
5	5.		(	3.	
x	-40.00	result of Mult.	2	2.	
ENTER	-40.00		-	2.	Subt. pending
6	6.		6	6.	
ENTER	6.00		)	-4.	result of Subt.
22	22.		)	-12.	result of Mult.
$\sqrt{x}$	4.690	SQ. RT.	)	-8.	result of Add
x	28.14	result of Mult.	=	-40.	result of Mult.
17	17.		÷(	-40.	defines divisor
÷	1.66	result of Div.	6	6.	
37	37.		x	6.	Mult. pending

REVERSE POLISH NOTATION			ALGEBRAIC NOTATION		
KEY	DISPLAY	COMMENT	KEY	DISPLAY	COMMENT
SIN	0.6018	results:SIN	22	22.	
+	2.26	result of Add	$\sqrt{x}$	4.69	results of SQ.RT.
÷	-17.7206	result of final Division	÷	28.14	results of Mult.
	-17.7206	ANSWER	17	17.	
			+	1.66	result of Div.
			37	37.	
			SIN	0.60	result of SIN
			)	2.26	
			=	-17.7206	results of final Division

Figure 1. A Test Problem as It Would be Entered in Reverse Polish Notation and in Algebraic Notation.

The test procedure required two groups of 24 subjects each, with one group for each calculator. Each group was further divided into two subgroups of 12 subjects each; one subgroup to solve the four problems without interruption and the other subgroup to solve each problem with a standardized 1 minute interruption beginning 12 seconds into each problem. During the 1 minute interruption, subjects were asked to write the multiplication table of nines, eights, etc. After 1 minute elapsed, subjects were asked to stop multiplying and resume solving the test problems. All subjects had paper and pencil available at all times.

A short explanation of the experiment was read to each subject and questions answered. The subject was given the test calculator, paper, pencil and the four problems, each on a separate paper placed face down. Subject was instructed to turn over one test paper on GO and proceed to solve the problem. The Interruption subjects were interrupted 12 seconds into the problem by "INTERRUPT" to perform the multiplications for one minute. At RESUME, these subjects returned to the problem signalling its completion by saying "FINISHED". The noninterrupted subjects proceeded from GO to FINISHED without interruption.

Four times were recorded by stopwatch. T_0 = GO, T_1 = first key pressed, T_2 = first key pressed after RESUME, T_3 = FINISHED. For the noninterrupted groups, only T_0, T_1, T_3 were recorded. Problem answers were recorded in all cases.

The 48 subjects had each been using a calculator for more than five years, most frequently on a daily basis. The subject population comprised undergraduate and graduate engineering students (60%) and faculty members (40%) of Engineering, Mathematics and Physics and one engineer from Hewlett-Packard. There were 27 males and 21 females ranging in age from 19 to 35 years. Subjects for each group (RPN, AN) were chosen from those owning and using a calculator with the respective notation system although all considered the particular test calculator as somewhat unfamiliar from the one they owned. Subjects were allowed to practice with the test calculator. Approximately 40% of the RPN subjects and 25% of the

AN subjects were faculty members although the faculty members reported using their calculators less frequently than did the nonfaculty subjects.

RESULTS

In comparing the two types of calculators the following six measures were computed.

(1) T_U - Total (uninterrupted) time to perform the problem

(2) T_I - Total time to perform the problem when interrupted (interruption time of 1 minute removed)

(3) T_N - Net effect on solution due to the interruption = $\bar{T}_I - \bar{T}_U$

(4) T_1 - Time to press first key

(5) T_2' - Time to press first key after the interruption ceases (T_2 - time at RESUME)

(6) E - Incorrect solutions.

Table 1 presents average and standard deviation values of the first three measures for each of the four problems as performed on the Reverse Polish (RPN) and Algebraic (AN) calculators. The fractional and unbiased percentage differences in solution times of each calculator are also shown.

The measures are computed for the 12 subjects in each of the four test sections.

TABLE 1

Interrupted An Uninterrupted Problem Entry Times For A Reverse Polish and An Algebraic Notation Hand Calculator.

	Avg (Sec) / Stnd Dev. (Sec)	Problem 1	Problem 2	Problem 3	Problem 4
Reverse Polish Notation	$\bar{T}_{I_{RP}}$	84.40 / 47.54	57.45 / 32.01	72.40 / 46.41	111.45 / 44.21
	$\bar{T}_{U_{RP}}$	77.58 / 34.06	45.67 / 25.32	69.85 / 38.02	110.35 / 23.13
	$\bar{T}_{N_{RP}}$	6.65 / 16.90	12.05 / 11.78	2.55 / 14.41	1.10 / 17.12
Algebraic Notation	$\bar{T}_{I_A}$	117.15 / 60.20	89.60 / 31.60	83.35 / 39.31	122.20 / 37.54
	$\bar{T}_{U_A}$	88.15 / 39.04	45.10 / 21.90	69.95 / 40.74	90.85 / 26.10
	$\bar{T}_{N_A}$	29.00 / 20.74	44.50 / 11.05	13.40 / 13.24	31.35 / 13.37
$\frac{\bar{T}_{N_A} - \bar{T}_{N_{RP}}}{\bar{T}_{N_{RP}}}$		3.36	2.69	4.25	27.55
$200\times\frac{\bar{T}_{N_A} - \bar{T}_{N_{RP}}}{\bar{T}_{N_A} + \bar{T}_{N_{RP}}}$		125 %	114 %	136 %	186 %

Figure 1 presents just T_U, $\bar{T}_I$ and $\bar{T}_N$ from Table 1 for both calculators. The test method allows only one sample of T_N to be computer for each problem in each group. However, the variance of T_N was computer as the sum of the two independent sample mean (T_U, T_I) variances and the corresponding standard deviation is shown in Table 1 for each T_N value.

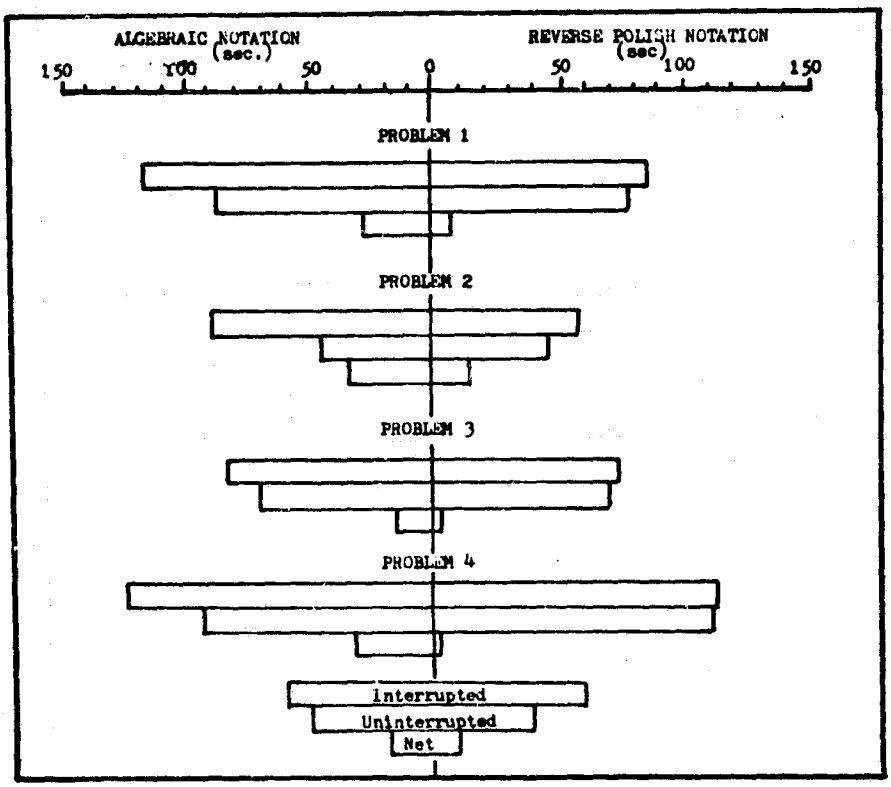


Figure 1. Interrupted and Uninterrupted Problem Entry Times for Reverse Polish and Algebraic Notation.

Both Table 1 and Figure 1 show that there was a consistently greater lengthening of solution time (T_N) on each problem for the Algebraic Notation compared to the Reverse Polish calculator when their operators were interrupted at the task. This difference was statistically confirmed by a student t test using the T_N value for each problem as sample data points and comparing the four values from the RPN with those of the AN calculator. ($t = 5.037$, $p < .02$). Table 1 shows that this statistically significant finding also represents a large unbiased percentage difference of at least 114% for the two calculators.

A similar student t test showed that the uninterrupted times did not differ significantly between the two calculators which is in agreement with several previous findings but contrary to Agate & Drury. Although the interrupted solution times (T_I) were consistenly longer on each problem for the Algebraic compared to the Reverse Polish Notation calculator, this difference

was not significant with the exception of problem number 2 ($t = 2.47$, $p < .05$).

The data were also reduced to obtain the time required before pressing the first key at problem start (T_1) and again following the end of the interruption (T_2'). The latencies are shown in Table 2.

TABLE 2

First Key Press Latency Times At Problem Start And Following the Interruption.

		PROBLEM				
		1	2	3	4	AVERAGE
START	$\bar{T}_{RP_1}$	8.15	4.2	6.47	7.45	6.57
	$\bar{T}_{A_1}$	7.15	3.6	8.65	6.18	6.40
INTERRUPTED	$\bar{T}'_{RP_2}$	4.80	4.35	5.05	5.29	4.82
	$\bar{T}'_{A_2}$	3.45	2.20	4.95	5.55	4.04

The only consistent finding among these latencies is that resuming the entry task after interruption required less latency than when first beginning it and the addition problem required the least latency of the four problems. There was no significant difference between the initial or between the post interruption latencies for the two calculators. It might have been assumed that the greater lengthening of entry time under interruption for AN compared to RPN would correspond to a similarly greater post interruption latency. However, this is not the case. If anything, the AN latencies appear a bit shorter (not significantly) rather than longer compared to those for RPN.

Incorrect solutions were also compared for the two calculators on each problem interrupted and uninterrupted and the results are set out in Table 3.

Although problem two showed the fewest incorrect solutions out of 12 (except for the AN - interrupted) only the Reverse Polish calculator - uninterrupted) showed a statistically significant difference between the problems ($x^2 = 7.33$, $p = < .06$).

TABLE 3

Number of Incorrect Solutions (Out of 12 Possible/Problem) For The RPN And AN Calculators Working With And Without Interruption.

		PROBLEM				
		1	2	3	4	AVERAGE
INTERRUPT	RPN	4	2	5	8	4.75
	AN	4	5	9	8	6.5
NON INTERRUPT	RPN	4	1	4	9	4.5
	AN	7	1	7	7	5.5

DISCUSSION

Previous experiments were equivocal whether Reverse Polish Notation had an advantage over Algebraic Notations regardless of problem complexity. The no difference findings were reconfirmed in the present experiment when sugjects were not interrupted during problem entry. When subjects were interrupted, as might naturally happen in real life the problem entry time increased for both notation systems (calculators) over that required under no interruption. However, the average increase for the Algebraic Notation calculator was more than twice as great as that for Reverse Polish Notation for the addition problem at the least and twenty-seven times as great at the most, on one of the complex problems. Regardless of the problem, the increase in entry time was longer for the AN calculator. Coincidently, although not statistically significant, more incorrect solutions occurred with the AN calculator as well.

Although it is tempting to attribute the difference in entry times to the nature of logic or notation it is difficult to support this conclusion directly. If one or the other notation system is per se more compatible with the subjects' own mode of problem organization, difference might be expected in the respective times taken to press the first key for each calculator either at the start of the problem or resuming after interruption. Presumably, these times might reflect the amount of "mismatch" between human and machine "logic" systems. However, these respective times were nearly identical for each calculator. (Again, and although not statistically significant, the addition problem required the least amount of this latency time which was also consistently less following the interruption than when just beginning the problem.) The negligible differences between key latencies, and between uninterrupted solution times suggests that the user can adapt to either notation system.

It is, in fact, difficult to find the mechanism of the entry time lengthening (T_N) in the interrupted mode. Neither the recording nor observational methods were sufficiently detailed enough to monitor each of the

individual key entry times. Examination of such a record might have suggested plausible mechanisms for these lengthenings. At present, and for lack of more detailed data, it could be assumed that any lengthening (T_N) is manifested as a uniform slowing of entry, for whatever reason, after interruption. It may be that problem recapitulation is necessary after the interruption and that the AN system places more memory stress on the user resulting in a longer time to execute the problem. With the increased memory burden adding to the task difficulty, resources invested in doing the entry task can do less, decreasing user efficiency (Novon and Gopher, 1979).

However, any mismatch between the human logic system, and the AN (or RPN) calculators thus can also be sought in key board and/or display differences.

A potential explanation of the solution time differences under interruption involves display rather than key board differences. The simple addition problem of [5 + 2 = (7)] as entered and displayed on each calculator is instructive for this purpose as shown in Figure 2.

RPN (HP)			AN (TI)		
Key	Display	Comment	Key	Display	Comment
5	5.		5	5.	
ENTER	5.00	'.00' shows that the number has been entered	+	5.	no indication to show if + has been keyed.
2	2.		2	2.	nothing to indicate 2 is not a resultant.
+	7.00	'.00' indicates 7 is not an entry but must be a resultant	=	7.	does not show if 7 is an entry or an answer.

Figure 2. Sample Entries of the Same Problem on HP(RPN) and TI(AN) Calculators.

As presently constructed, the HP (RPN) calculator display differentiates between resultants and entries while the TI (AN) calculator does not. Of course, if the entry had exactly as many decimal places as the displayed resultant, this difference would no longer hold. A second display difference appears in the familiar left-to-right concatenation of digits as they are entered on the HP mimicking the normal mode of writing as opposed to the left shift of the entire digit string as a new numeral is entered on the TI calculator.

The HP display features may be a better aid to "place holding" during interruption than those of the TI calculator permitting a more rapid recovery as the problem is resumed.

At present it is not unequivocably possible to attribute the large percentage differences in interrupted solution times to notational or display differences or both for the two calculators.

However, from a practical point of view, the HP calculator (RPN notation and display) is superior in terms of the shorter time taken to complete a problem when interrupted - a common real life occurrence - and would be recommended where many such computations must be made or where time is a critical factor.

Since task interruption is a fact of life, and since it also degrades performance, it is reasonable that the logic/display design be as performance resistant to the interruption as possible. It is suggested that on computer systems without printers, the display show simultaneously at all times the last datum entry and operation as well as the current resultant and that a review key permit sequential viewing of the previous few datum + operation pairs as though a key stroke list had been internally compiled for later display. Although no definite suggestion can be made for the notational system itself, it is possible that some hybred of the RPN and AN may be more compatible with the user's own modes of problem phrasing and recall.

Further studies will be made of the notational and display features unconfounded in search of the "ideal" calculator which at this point could be said to produce identical interrupted and uninterrupted solution times regardless of problem complexity in addition to other desirable features. For design and evaluation purposes, it is suggested that interrupted solution times serve as one of the performance criteria.

REFERENCES

(1) Agate, S., Drury, C., Electronic Calculators - Which Notation is Better? (1980) Applied Ergonomics 11 (1), March.

(2) Card, Stuart, K., A Method for Calculating Performance Times for Users of Interactive Computing Systems. (1979) Proc. 1979 International Conference on Cybernetics and Society. Denver.

(3) Chapanis, A., Research Techniques in Human Engineering (1959) Johns Hopkins Press.

(4) Durding, B., Becker, C., Gould, J., Data Organization (1977). Journal of the Human Factors Society, 19 (1), 1-14.

(5) Enstrom, K., Rouse, W., Real Time Determination of How a Human Has Allocated His Attention Between Conrol and Monitoring Control. (1977) IEEE Transactions on Systems, Man and Cybernetics, SMC-7 No. 3.

(6) Grace, G. L., Preface to Human Factors Journal Special Issue on Information Processing Systems. (1970) Journal of the Human Factors Society. 12(2), 161-164.

(7) Kasprzyk, D., Drury, C., Bialas, W., Human Behavior and Performance in Calculator Use With Algebraic and Reverse Polish Notation, (1979), Ergonomics, 22(9) 1011-1019.

(8) Novon, D., Groper, D. Interpretation of Task Difficulty in Terms of Resources, Efficiency, Load, Demand, Cost, Composition. (1979) Technion Institute, Haifa, Israel.

A MODEL FOR THE CONTROL MODE MAN-COMPUTER INTERFACE DIALOGUE

By Roy L. Chafin

Jet Propulsion Laboratory
California Institute of Technology

SUMMARY

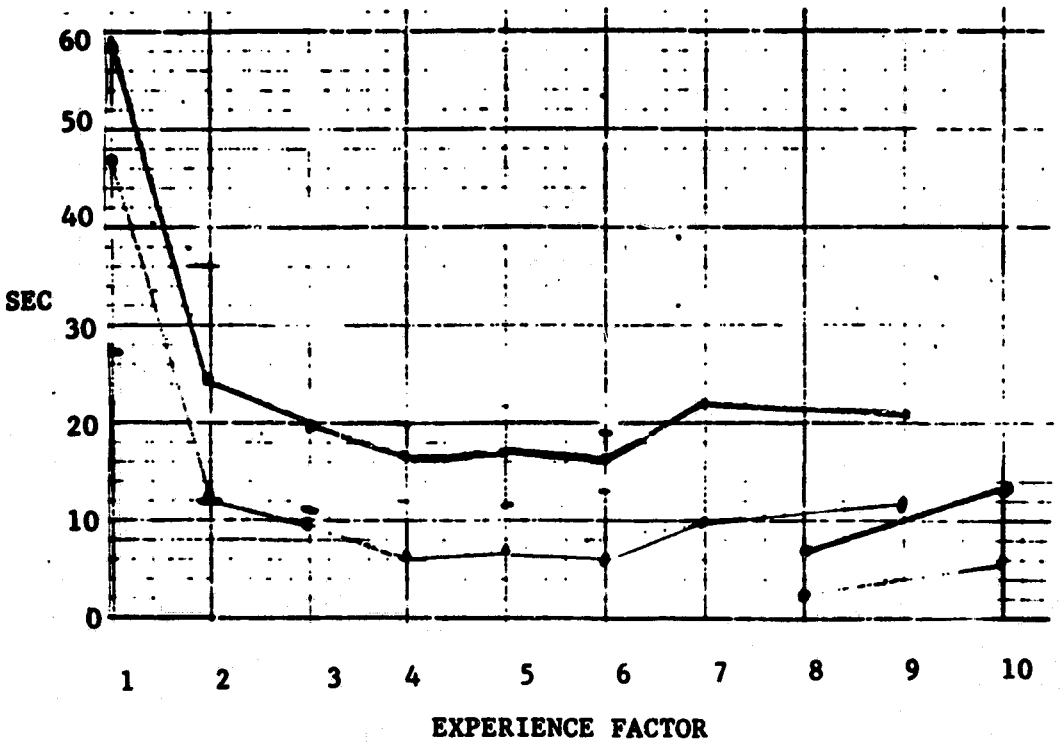
A four stage model is presented for the control mode man-computer interface dialogue. It consists of context development, semantic development, syntactic development, and command execution. Each stage is discussed in terms of the operator skill levels (naive, novice, competent, and expert) and pertinent human factors issues. These issues are human problem solving, human memory, and schemata. The execution stage is discussed in terms of the operators typing skills. This model provides an understanding of the human process in command mode activity for computer systems and a foundation for relating system characateristics to operator characteristics.

INTRODUCTION

Computer systems have two basic modes of operation, the control mode and the data mode. In the control mode, the operator controls the system by commanding it to take specific actions. For a telemetry system, it might be to acquire a specific data stream. For a teleoperator system, it might be to extend the arm and pick up an object. For a text editor system, it might be to delete some portion of the text or to place the text in a specified file. In the data mode, the operator is either entering data into the system or retrieving data from the system. For example, after a text editor has been commanded to accept text for insertion into a specific location, the text to be inserted is entered. That is the data entry mode. Or for a Data Base Management system, a data request is entered in the control mode and the data is presented to the operator in the data retrieval mode. This paper is concerned with only the control mode.

The concepts discussed in this paper are the result of reflections on data taken from a human factors experiment performed in the Deep Space Network (DSN) at the Jet Propulsion Laboratory, a NASA facility (1). The experiment was a man-computer interface test with approx. 100 operators from the DSN. The subjects were given a series of tasks on a CRT display of a simulation computer. They had been randomly assigned one of four command formats, single argument mnemonic, multiple argument mnemonic, prompt, or menu. They entered the command format into the keyboard to accomplish the task. Their solution (the command) was displayed on the CRT for feedback. It was also timed and recorded on disc for subsequent data reduction. Fig. 1 illustrates the performance time of one of the formats for a specific type of task.

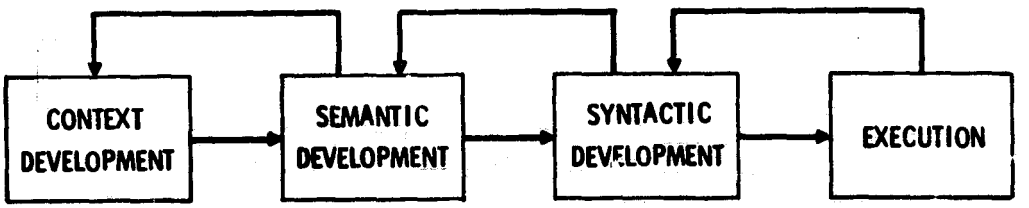
668



Experimental Performance Time

Fig. 1

The vertical dimension is performance time, the time required to complete the task. The horizontal dimension is the number of times this specific type of task has been attempted, it is an experience factor for the specific command format and the specific type of task. The first time this type of task was attempted required an average of 59 sec to complete the task. For the sixth attempt, the average task completion time was 16 sec. The first six attempts were consecutive. The seventh and nineth attempts were separated from tasks of the same type by a number of different type tasks. The eight and tenth attempts were partial tasks and are not of interest in this discussion. What is interesting is the increased time required to complete a task when it has been separated from previous tasks of the same type (ie. 6th at 16 sec and 7th at 22 sec). At least two explanations can be offered for this performance differential. One is that the subjects have forgotten because of intervening tasks. The other is that each task requires a context to be developed and subsequent same task allows the subject to keep the preceding context. Intervening tasks require that the subject change the context and that requires some time. The lower curve is the time required to enter the first character of the command. It represents the time required to compose the command, that is the think time. It produces the principle variance in the overall performance time. The time required to actually execute or type in the command is the time between the two curves. A four stage model (Fig. 2) is proposed to represent the total time required to generate or compose a command in the control mode. It can be used to explain the experimental performance time as typified in Fig. 1.



Control Mode Operator Model

Fig. 2

The first stage is the context development, the operator's definition of the domain of relevant information for the specific tasks being addressed. The second stage is the semantic development, the understanding of the factors and relationships which apply to the command generation. The third stage is the syntactic development of the command, the actual codes and symbols which make up the command. The fourth and last stage is the execution of the command, typing it into the keyboard and verifying its operation.

HUMAN FACTORS ISSUES

Several human factors issues interact in the control mode model. A very important issue is operator skill level. At some skill levels the command development is basically a problem solving exercise. At other levels it is basically a memory exercise. At some skill levels, the command development is a cognitive process and at other levels it is an automatic process(schema). This section presents an operator taxonomy based on skill levels and discusses human problem solving, memory, and schemata.

Operator Taxonomy

We intuitively understand that operators do not all have the same capabilities and skills, however much of the literature and most applications do not take operator variablility into account.

Operators vary over many dimensions. Eason (2) uses a "kind of user" taxonomy of clerical, manager, and specialists. Clerical users are principally data entry operators. Managers are principally data retrievers. And specialists use computer systems as a tool to accomplish some specific job. Bennett (3) divides users into those who are committed by their jobs to using the computer system and those whose computer use is discretionary. Similarly, Codd (4) divides users into those casual users who infrequently use the system and those dedicated users who frequently use the system. We would expect the manner in which they most effectively use the system to be different.

These taxonomies are related to how the user makes use of the system. Another interesting dimension is skill levels. Eason (5) also considers naive users who use the system as a tool but that do not have a deep knowledge of

the system. He implies that they are relatively unskilled. Hiltz and Turoff (6) suggest a four phase user skill development from their experience with computer conferencing systems. The user initially approaches a system with uncertainty. He progresses to a stage of insight when he understands the general concepts of the system. The incorporation phase is when the mechanics of the system interaction become second nature, a part of the users normal environment. And saturation occurs at some point in their experience.

The Hiltz and Turoff four phase skill development taxonomy can be generalized by considering that these four phases or stages can be static as well as dynamic. If a user is a casual operator, he may never develop beyond the insight stage. The skill level may be a function of the kind of system and the application tasks as well as a transitory development phase. To provide a generalized operator skill taxonomy, skill levels will be defined for naive, novice, competent, and expert operators.

Naive operators are those who have essentially no understanding of the system. They must rely on external assistance (either other users, trainers, or documentation) in order to use the system.

Novice operators are those who have a general but not a specific understanding of the system. They know what the system does but typically not how to operate it. They still need external assistance but of a different kind. They need a demonstration of how to operate the system.

Competent operators are those who understand the system and can use it effectively. Their knowledge of the system is sufficient for them to determine the actions required to control the system, primarily a cognitive process. They do not require external assistance beyond possibly an occasional reference to the user manual to refresh their memory.

Expert operators are those who know the system so well that they do not have to think about the control actions, their actions are automatic.

Problem Solving

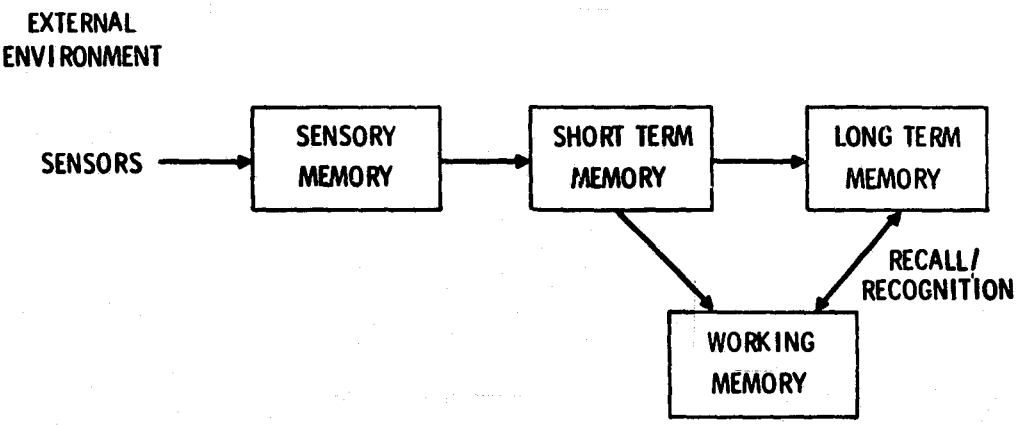
Problem solving is the process of creating a solution to a given problem. Over past years there have been many problem solving models presented (7). They tend to represent the originators perception of the process and the specific types of problems being solved. The problem solving models range from 4 to 9 stages. Osborn (8) suggests a 7 stage generalized model from which we will select a 5 stage problem solving model that is appropriate for the development of the various stages of the control mode model.

* Preparation - gathering the pertinent data.
* Analysis - breaking down the relevant material.
* Hypothesis - piling up alternatives by ways of ideas.
* Synthesis - putting the pieces together.
* Verification - judging the resultant ideas.

The pertinent data is gathered from the problem statement, the documentation, and from the operator's previous experience on similar problems. The analysis identifies the relevant elements so that the relationships between them can be evaluated. A number of the most reasonable alternative relationships are selected for further consideration. Synthesis puts the elements together within the relationships hypothesized, and selects the one best solution. That solution is evaluated against the problem statement (or task) and the documentation (or the remembered items from the documentation). Any discrepancy causes iteration back to previous stages in the problem solving process.

Memory

Shniederman (9) refers to a 4 element human memory model. He used the model in terms of defining the programming process but it is also very useful in the development or generation of commands in the man-computer dialogue. The model is presented in Fig. 3.



Human Memory Model

Fig. 3

The external environment is seen, heard, felt, etc., through sensors into the sensory memory. Information is stored in the sensory memory for a very short time, a matter of only fractions of a second. For example, an image of a printed page would be stored in the sensory memory. A part of the information in the sensory memory is passed onto the short term memory where it is held for a few seconds. For example, a set of characters will be selected from the page image in the sensory memory, interpreted, and stored in the short term memory. Short term memory seems to fade significantly after some 5 to 20 seconds, unless it is rebuilt by a process called rehearsal. To hold information in short term memory for longer periods of time, the individual concentrates or reiterates that information, that is, rehearses it. In addition to time limitations, it is also capacity limited. Short term

memory is limited to about seven units of information (10). Many of the operator errors that are experienced in system operations are due to system demands which exceed the individual's short term memory capability. Examples of this problem are the commands that require 8 to 12 digit frequency numbers (example, 2202786.012 Hz). A method of handling the bigger numbers or groups of characters is available by grouping the characters into meaningful combinations, each of which is easy to accomodate. This is called "chunking". It is illustrated by the common representation of telephone numbers. Even though there may be up to 15 digits in a telephone number (area code,number, & extension) it is grouped into 3 or 4 digit chunks, each with a specific meaning (area code, exchange, line or instrument, and extension number).

Again, some part of the information in the short term memory is passed to the long term memory where it is held permanently. There seems to be no limit on the amount of information stored in long term memory. Two retrieval methods from long term memory have been suggested (11). One is recall, in which information is recalled directly from memory. The other is recognition, in which the information that cannot be obtained from direct recall can be recognized when matched with some external sensing. For example, an operator can recall a command mnemonic for a command that is used frequently, but cannot recall (or remember) the mnemonic for a command which is used infrequently. However, he can scan a list of mnemonics and readily recognize the correct one when he sees it.

Working memory is the area used in processing the task or problem. In the context of generating commands, it is the area where the problem solving activity occurs. The working memory receives information from the external environment through the sensory and the short term memories and from the long term memory directly. Parts of the command development process is stored back into long term memory to be used in subsequent command generation activities.

Schemata

Another important concept is the automatic actions of an expert operator doing a repetitive routine task. It is called a schema. As Zipf (14) has suggested in his studies on the use of language, people tend to use the minimum effort in accomplishing tasks. When a task has become repetitive and routine an individual develops a scenario or schema which he can use to accomplish the task without thinking about it (15). Once the task has been identified and the proper schema triggered, the individual goes through the actions automatically. He doesn't have to consider and think through each action in the schema, and his decision requirements are minimized. An example that is familiar to all of us is driving to work. After driving the same route for several years, we can drive it without thinking. We do not have to decide where to turn, how fast to go, where to slow down, how fast to take the curves, etc., we do all these things automatically.

When an expert operator on a system uses schemata, his work load is decreased and his performance is increased. Of course, we have to be careful that the basic action sequence does not change without our realizing it. That

would make the schema inappropriate and would lead to errors and incorrect results. A common cause of operator errors with "He is one of our best operators" is due to the operator triggering a schema without realizing that the task has subtly changed. System operating procedures can be error prone when the system does not provide adequate clues to the operator that the task is different from his routine task.

OPERATOR MODEL

The command mode operator model consists of the context development stage, the semantic development stage, the syntactic development stage, and the execution stage.

Context Development

Context is an essential element in human discourse. Without a mutual understanding of some context our speech would be hopelessly long, involved and difficult to follow. As human communicators, we typically assume a context based on our understanding of the other person, the situation, and past history. This usually works, however, a more interesting situation is when we do not have these clues and we have to use our skills to develop the context under which we will carry the conversation. Our conversational success then depends upon these skills. An advantage in context development between humans over man-computer context development is the flexibility that both sides of the human conversation can bring to the process. The typical computer system is very constrained in this issue. Although there are some exotic programs coming out of the artificial intelligence field in which the program participates in the context development (16), the programs which are developed for most applications demand that the operator develop the context.

Grosz (17) defines a domain of discourse. Without the ability to focus on the subset of knowledge relevant to a particular situation, the amount of knowledge overwhelms even the simplest system. The process of defining this domain of the discourse and to limit the knowledge base needed for a particular application is what we will define as context development.

Context development is the successive narrowing of focus from the general to the specific. Context development then is a selection process. It uses either problem solving or memory depending on the skill level of the operator. Naive operators have no experience to provide memory capability, so context development is very much a problem solving process. The system characteristics which aid this process are a well structured man-computer interface (MCI) design and knowledge aids. The MCI structure is most effective when it requires a minimum of selection at any hierarchical level, and each selection is well identified. Understanding must be developed at each selection, therefore, the system documentation becomes extremely important. As the

operator's skill increases, he operates more on understanding and memory. At the novice level we would not expect the memory to be extensive, it would most likely give clues to the problem solving process. Because the understanding is somewhat greater, the external documentation is of less value. The more important MCI characteristic is cognitive simplicity (18). Cognitive simplicity is the use of internal MCI aids to help the novice operator's understanding.

The competent operator has the experience that allows him to use his memory for the context selection process rather than going through a problem solving process. The expert operator has developed schemata to accomplish the context development. All he requires is to identify the situation and he will trigger the appropriate schema. The competent and expert operators require an environment which is very straight forward. They would prefer to go right to the context rather than going through a series of levels or stages. This requires a different MCI organization than for naive or novice operators.

Normal system operation requires a change of context or focus as the system sequences through its tasks. As the task which the system has to accomplish changes, the context of the MCI also has to change. This situation is very similar to the starting operation context development with two additional steps involved. The first step is recognition, the operator has to recognize that the task has changed and that the context must then change. The second is evaluation, he must evaluate where he is and where he has to go. For naive and novice operators with highly structured MCI's the process must work in reverse until they navigate back to a level which allow them to go forward again. An MCI design which caters to naive and novice operators must take particular care to provide for this need. MCI's designed for competent and expert operators do not have this problem, they can and prefer to go directly to the new context.

Semantic Development

After the context has been established, the next stage is semantic development. Command semantics is the knowledge of how the command relates to the task that it is supposed to accomplish. A command consists of a function select and possible arguments to satisfy the required parameters for that function. Some examples:

1. PUMP
2. PUMP/1,ON

The semantic knowledge associated with "PUMP" in #1 is that it controls the pump and turns it on. The "PUMP" in #2 refers to more than one pump and an argument is required to select the desired pump (ie. 1). Also, the function select "PUMP" in #2 can turn the selected pump on or off, so an argument is required to determine whether the pump is to be turned on or turned off. This semantic knowledge is independent of the command style, that is, whether it is mnemonic, prompt, or menu style. This is the semantic development required of the operator, he must understand that part of the system that he is attempting to control.

The naive operator must use a problem solving process for semantic development. Without experience, he has nothing to go on. Three things will help the naive operator to develop the semantic understanding. One is good documentation. That is, documentation or user manuals that allows the operator to rapidly find the function (context) and explictly defines the function select and the arguments. This would not likely be the Theory of Operations section of the manual because it tends to be too general and too bulky. It would most likely be the Operator Instructions section of the manual because it tends to be more direct. Second, the system can be designed to have compatibility with the operator's previous experience. Compatibility is a technical term in human factors which means that a process is what a person expects it to be. His expectations may be due to past experience with similar systems or to a more natural connection such as a car steering wheel turning right for a right turn. Compatibility is a powerful way to help the operator's semantic development. Third, the MCI can be designed to be "User Friendly", which seems to be a buzz word for a menu driven system. Menu systems tend to help the operator in context development because it leads him through the choices. It is helpful in the semantic development if it is sufficiently explicit, however, this tends to produce menus which are very wordy. Another characteristic of a menu system that is helpful to a naive operator is that, at any level, all the choices are available to the operator. He may not have to understand the function completely if he is able to correctly differeniate between the choices (like guessing on a multiple choice test). But again, the menus must be explict or a high error rate will be experienced.

A novice operator has more experience to draw on. His semantic development will most likely be a combination of memory recall and documentation referal. Even when using the documentation, he will very likely be using the recognition memory mode, he will scan the manual and recognize the command when he sees it. Prompt and menu MCI formats are appropriate at this skill level. The menus can be less explict and less wordy at this level. In fact, they should be less wordy or they will become unattractive.

The competent operator tends to work primarily from the recall memory mode. And the expert tends to operate from his schemata. For both skill levels, the so called "User Friendly" MCI's are not really friendly. They tend to be too long and involved to be comfortable. These operators tend to lose patience with prompts and menus because their own pacing is faster than the pacing of the MCI.

Syntactic Development

Shneiderman (9) points out that syntactic knowledge is the second kind of information stored in long term memory. He also points out that it is more precise, detailed, and arbitrary, he also suggests that it is more easily forgotten. Sachs (19) supports this suggestion from work in recognition memory for syntactic and semantic aspects of sentences. The meaning of sentences is much easier to remember than the exact syntax. Of course, in human discussion the meaning is important not the exact syntax. Although, philo-

sophically, the same comment applies to man-computer communications, ie. it is the meaning that is important, practically the limitations on the "understanding" capability of the computer increases the inportance of exact syntax considerably.

In an MCI, the syntax is the specific codes and symbols used to specify the command and its arguments, the punctuations or delimiters, and the structure that ties these elements together. The naive operator must determine the syntax from the documentation or from the display for prompts and menus. His task is basically problem solving. The problem is to determine the proper syntactic elements to implement the semantic development. Aids to the novice's syntactic development are things that tie into the semantic content or are compatible with prior experience. This is the attraction for so called natural language MCIs, they are supposedly compatible with human communication syntax. Menus are appropriate for naive operators because they minimize the syntactic elements that he has to create, he only has to choose between the elements presented to him.

By definition, a novice operator has been exposed to the syntax of an MCI. He would most likely operate in a problem solving mode for some syntactic elements, from recognition memory for others, and he may be able to recall other elements. And as he progresses in experience he develops the capability to recall more of the syntactic domain. The aids which are important to the naive operator are still useful to the novice operator. He will most likely refer to the documentation and would respond favorablely to menus but he is more able to operate independent of these aids. The documentation that he uses would most likely be the Operator Instructions rather than the Theory of Operations sections and as he becomes more proficient he would prefer quick-look MCI tables. Cognitive simplicity (18) becomes more important for the novice operator because he is often operating from the recognition memory mode.

The competent operator will be working from the recall memory mode and as such doesn't need the recognition memory aids or the problem solving aids which are so useful to the naive and novice operators. He will use the documentation infrequently. He would prefer process simplicity (18) because it is easier to execute. Process simplicity is the concept of minimizing the execution effort. Prompting and menus may be tolerated but they are inappropriate and he would prefer the more straight forward mnemonic command style. He would prefer the increased control which he has with the menemonic command formats.

The expert operator will be working from schemata. The problem solving and memory aids are inappropriate for him, in fact they are undesirable because they interfere with schema development. They tend to make the schema longer, more involved, and harder to execute. The expert operator is likely to actively dislike a prompt or menu command format because they get in his way (3). They would prefer extreme process simplicity.

A complicating factor is that very often an individual operator's skill varies greatly over the entire syntactic domain of a system. He may be an expert in the commands and functions that he uses often, and naive in the others that he doesn't use so often. The challenge for the system designer on an extensive system is to gracefully handle this extreme skill range.

Execution

The execution phase of the model is the actual execution of the command. Card's (20) keystroke model was for expert users performing routine tasks. This model is concered only with the command entry and does not include any data or text entry. Card provided an excellent discussion of the elements of the command execution and this paper will not discuss these items further. A generalized command mode model must include a broad range of tasks and operator skills. Once the operator has developed the syntactic representation of the command, the skill differential is mainly in typing capability. Operators can be skilled typists or unskilled typists. A task which contains considerable data or text entry usually demands a skilled typist but one which is primarily command entry only can use either skilled or unskilled typists. The principle difference between skilled and unskilled typists is the use of the touch system for the skilled versus a hunt and peck system for the unskilled typists. Another difference is in the need to look at the keyboard when typing. The skilled touch typist can maintain his attention on the task while entering the command, he doesn't have to divert his attention from the task to look at the keyboard. The unskilled typist must take his attention away from the task and apply it to selecting the keys on the keyboard. Changing attention requires the use of short term memory to hold the information. The unskilled typist operator's performance is influenced by his short term memory limitations whereas the skilled typists are not operating under this limitation. The short term memory limitation causes the operator to have to chunk information into small groups, a process which can be error prone if the MCI was not designed to accomodate it.

Task and typing skill interact to influence the operator performance. A task which demands the operators continous attention will suffer with an unskilled typist and should have a skilled typists. A task which allows the operator to compose the command syntax in his head and does not require continous attention would be appropriate for either a skilled or unskilled typist operator.

Some observations from the typing tests of the DSN Human Factors experiment (1):

1. The performance difference between trained and untrained typists is approximately 1:2.

2. Random characters are more difficult to type than English text. This supports other findings in the literature (21).

3. The trained typist's speed for a different number of random characters in a group is quite constant. This supports the position that they do not have to take their attention away from the task.

4. The untrained typists speed decreased for increased number of random characters in a group (over the range 3 to 9 characters per group). This supports the short term memory influence on untrained typing performance.

5. The above comments apply to alphabetic characters only. When the full range of ASCII characters were allowed in the random character groups, whether the operators were trained or untrained was immaterial. Apparently typing training doesn't handle numbers, punctuation symbols, or other special symbols very well.

This execution model which includes a variable for trained/untrained typists suggests that consideration for operators typing skill should be included in the MCI design. If the expected user population contains a high proportion of untrained typists, short term memory limitations should be seriously considered in the MCI design. It also suggests that, regardless of the typing skill, the MCI should be designed with commands that are familiar to the operator rather than what might seem like a collection of random characters.

CONCLUSION

A four stage model of the control mode command generation process has been presented. It consists of the subelements of the cognitive process of composing the command and physical element of the execution of the composed command. The composition subelements are the context development, the semantic development, and the syntactic development.

The value of this model is in the understanding of the human process that it gives to the system designer when designing the system's man-computer interface. It provides a foundation for relating MCI characteristics to different operator skill levels. We would expect that matching these characteristics (man and machine) would provide systems that are "easy to use", have few errors, and have better user satisfaction.

REFERENCES

1. Chafin, R.L.; and Martin, T.H.:DSN Human Factors Project Final Report, Jet Propulsion Laboratory, Calif. Inst. of Technology, 4800 OakGrove Drive, Pasadena, Calif.,Contract No. 955013, Task Order No. RD-142, 1 Nov. 1980.

2. Eason, K.D.; Damodaran, L.; and Stewart, T.F.M.:Interface Problems in Man-Computer Interaction, Human Choice and Computers eds E. Mumford and H. Sackman, North-Holland Publishing Company, 1975.

3. Bennett, J.L.: Incorporating usability into System Design: the Opportunity for Interactive Computer Graphics, Proceedings of the International Conference on Cybernetics and Society, Tokyo, Japan, Nov. 1978.

4. Codd, E.F.: Seven Steps to Rendezvous with the Casual User, Data Base Management eds J.W. Klimbe and K.L. Koffeman, North-Holland Publishing Company, 1974.

5. Eason, K.D.: Understanding the Naive Computer User, The Computer Journal, Vol. 19, No. 1, pp 3-7, Jan. 1976.

6. Hiltz, S.R.; and Turoff, M.: The Network Nation, Human Communication via Computer, Addison-Wesley Publishing Company, Inc., Advanced Book Program, Reading, Massachusetts, 1978.

7. Sackman, H.: Preliminary Investigation of Real-World Problem Solving with and without Computers, Volume II: Complete Results, R-1205/1-NSF, Rand Corp., Santa Monica, Calif. 90406, April 1973.

8. Osborn, A.: Applied Imagination, Charles Scribner's Sons, New York, 1954.

9. Shneiderman, B.: Software Psychology Human Factors in Computer and Information Systems, Winthrop Publishers, Inc., Cambridge, Massachusetts, 1980.

10. Miller, G.A.: The magical number seven, plus or minus two: Some limits on our capacity for processing information, Psychological Review, Vol. 63, pp 81-97, 1956.

11. Loftus, G.R.; and Loftus E.F.: Human Memory: The Processing of Information, New Jersey, Lawrence Erlbaum Associates, Publishers, 1976.

12. Cheriton, D.R.: Man-Machine Interface Design for Timeshare Systems, Proceedings of the Annual Conference, Association for Computing Machinery, New York, 1976.

13. Watson, R.W.: User Interface Design Issues for a Large Interactive System, Proceedings of the National Computer Conference, Vol.45, pp 357-364, 1976.

14. Zipf, G.K.: The Principle of Least Effort, Addison-Wesley Press, Inc., Cambridge, Massachusetts, 1949.

15. Oldfield, R.C.:Memory Mechanisms and the Theory of Schemata, The Cognitive Process Readings, Harper,R.J.C., et al. eds., Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1964.

16. Hendrix, G.G.: Human Engineering for Applied Natural Language Processing, Proceedings of the 5th International Joint Conference on Artificial Intelligence, pp 183-191, 1977.

17. Grosz, B.J.: The Representation and Use of Focus in a System for Understanding Dialog, Proceedings of the 5th International Joint Conference on Artificial Intelligence, pp 67-76, 1977.

18. Chafin, R.L.: Simplicity in Command and Control Systems: A Human Factors Considersion, Proceedings of the 1980 International Telemetering Conference, 14-16 Oct. 1980, San Diego, Calif., 1980.

19. Sachs, J.S.: Recognition Memory for Syntactic and Semantic aspects of Connected discourse, Perception and Psychophysics, Vol. 2, No. 9, pp 437-442, 1967.

20. Card, S.K.; Moran, T.P.; and Newell, A.: The Keystroke-Level Model for User Performance Time with Interactive Systems, Communications of the ACM, Vol. 23, No. 7, pp 396-410, July 1980.

21. Schoonard, J.W.; and Bores, S.J.: Short-Type: A Behavioral Analysis of Typing and Text Entry, Report RC 4434, July 16, 1973, IBM Thomas J. Watson Research Center, P. O. Box 218, Yorktown Heights, New York 10598.

D58

STRUCTURE OF THE KNOWLEDGE BASE FOR AN EXPERT LABELING SYSTEM*

By N. S. Rajaram

Lockheed Engineering and Management Services Company, Inc.

SUMMARY

One of the principal objectives of the NASA AgRISTARS program is the inventory of global crop resources using remotely sensed data gathered by Land Satellites (Landsat). A central problem in any such crop inventory procedure is the interpretation of Landsat images and identification of parts of each image which are covered by a particular crop of interest. This task of "labeling" is largely a manual one done by trained human analysts and consequently presents obstacles to the development of totally automated crop inventory systems. However, development in Knowledge Engineering as well as widespread availability of inexpensive hardware and software for Artificial Intelligence work offers possibilities for developing expert systems for labeling of crops. Such a knowledge based approach to labeling is presented in this paper.

INTRODUCTION

The Landsat spans different parts of the earth's surface providing images at regular periodic intervals. Images gathered by Landsat are in the form of picture elements (or pixels) consisting of the average spectral response of the area covered by each pixel (about an acre) in four visible and near infrared frequency bands. It is known however that it is possible to make a transformation of this four-channel data onto a two-dimensional plane essentially preserving the information contained in the original data. This permits creation of visual displays to be used by analyst interpreters (labelers) for assigning labels to different parts of the scene. (cf: Ref. [2], [4], and [6]). It is to be noted that labeling forms only a part of the overall remote sensing crop inventory exercise, but as experience has shown, a most crucial part. It may also be mentioned that labeling of images arises in other applications besides crop inventory by remote sensing (Ref. [9]).

In all decision making processes including labeling, human experts use knowledge which is not easy to formalize. Such expert knowledge can frequently consist of heuristic rules, elimination by constraints, exercise

*Work for this project being performed for the NASA Earth Resources Division under Contract NAS-9-15800 at the Lyndon B. Johnson Space Center.

of judgment, all gained through experience. In the context of labeling, experienced analysts make successful use of spatial context and texture. It is precisely these skills which one would like to impart to an automatic labeler.

One of the most intriguing concepts in automation technologies is the idea of employing Artificial Intelligence (AI) systems in a decision making capacity, in effect replace human experts by "Expert Systems." One of the most active areas of AI research is the application of such techniques to problems in Pattern Recognition/Image Understanding (Ref. [8] and [9]). This development is hardly a coincidence. AI and Pattern Recognition share several common features among which the most significant is that they are "knowledge" based (Ref. [8]). A fundamental requirement in AI approaches is that one be able to represent and process "knowledge" and not merely data. We briefly discuss a knowledge based approach to labeling in remainder of this paper.

KNOWLEDGE BASED APPROACHES

Knowledge Engineering is the branch of Artificial Intelligence used in building expert systems. In it one attempts to capture the essential problem solving skills of an expert, transmit those same skills to a computing system thereby creating an automated expert or an expert system. It is now recognized (Ref. [2]) that human problem solvers possess knowledge and techniques which are specific to a problem area and not general problem solving skills. Furthermore, this knowledge is frequently heuristic knowledge consisting of judgment, experience, good practice, and so on. Thus, it is clear that any expert system has to include a store of knowledge called "Knowledge Base" and a set of techniques usually called "Paradigms." Consequently, an effective and flexible representation of knowledge is a crucial first step in attempting to develop any expert system.

The pioneering effort in knowledge based problem solving was undoubtedly Slagle's integration program (Ref. [8]). Since that time many expert problem solvers ranging in applications from advanced mathematics to medical diagnosis have been successfully built. Of these we can cite MACSYMA (advanced mathematics), MYCIN (diagnosis of blood disorders), and DENDRAL (chemical analysis of spectroscopic data) as examples of highly successful knowledge based expert systems. As a result of these efforts, the following general insights have been gained.

1. Any Artificial Intelligence work entails storage and manipulation of complex data structures.

2. These complex data structures which represent "knowledge" can include data as well as programs. Hence, the programming environment must facilitate uniform representation of data and programs.

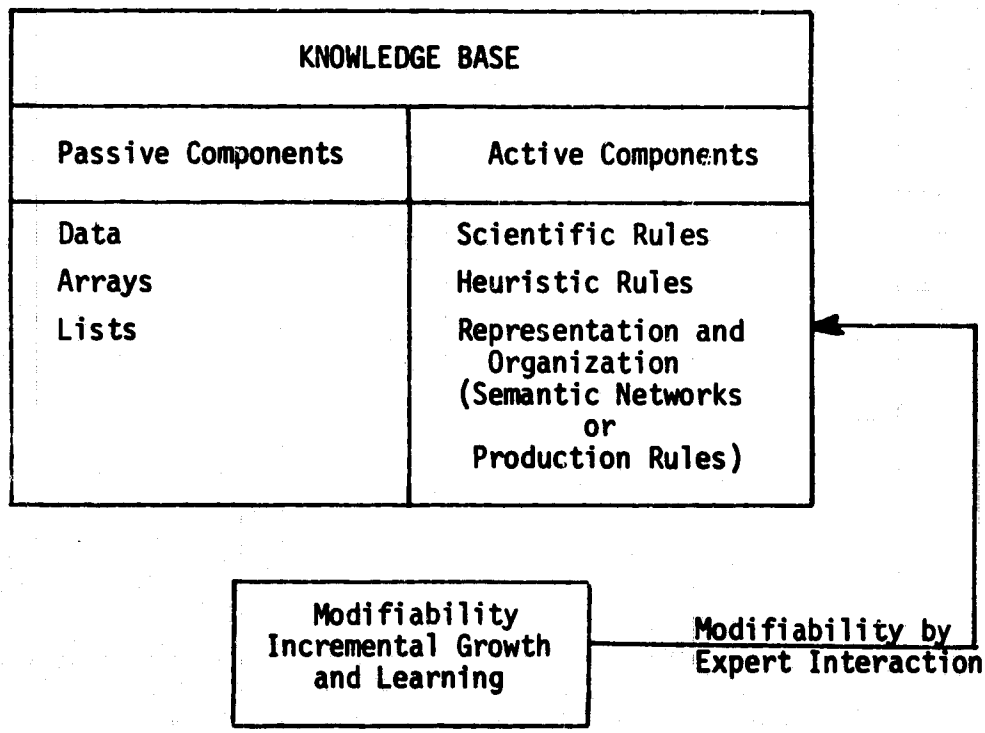
3. This activity involves "knowledge processing" and not just "data processing."

4. Such a knowledge based system should permit ease of frequent modification as we update our knowledge, without dismantling the system. This is known technically as "embedability."

5. These conditions argue for a LISP base for AI systems. In fact LISP is the sine qua non for serious AI work. (Ref. [8])

In addition to its uniform treatment of data and programs as symbolic expressions, LISP as a programming language provides several extremely powerful features such as lambdas and recursive function calls for transfer of control and use of programs and functions as arguments for other programs.

We can represent a knowledge base by the following general schematic diagram.



KNOWLEDGE BASE FOR THE EXPERT LABELING SYSTEM

Recall that we are interested in assigning different parts of a scene to different land cover types based on its Landsat imagery or "label" the scene. We also noted that the 4-dimensional data obtained by the multispectral scanner (MSS) are transformed into a two-dimensional space known as the "greenness-brightness" space. Consequently, data to be used for labeling will consist of multitemporal greenness-brightness values for each pixel. We shall see later that there are certain minimal data requirements. The production rules will be based on a study conducted by Palmer and Magness (Ref. [5]), and contain a very large heuristic base. These rules will be displayed in a widely used version of LISP known as "MACLISP" (Ref. [9]). We shall not discuss the labeling paradigms separately but merely note that inference rules with the help of IF THEN--- rules by COND constructs present no particular difficulties (Ref. [9]).

Because of atmospheric conditions not all the Landsat passes over a scene result in usable data acquisitions. Such deficiencies result in imperfect or incomplete labeling. For the sake of exposition, we assume that we are attempting a two-stage decision, assigning a label as Spring Small grains or not at stage 1 and as barley or not at the final stage. In order for a pixel to be labelable, the minimal acquision requirements are expressed in terms four "windows" which are time intervals in which data are available. Based on Palmer-Magness study, we define the four "windows" which are intervals represented by lists as follows:

```
(SETQ W1 (LIST

(DIFFERENCE X1 5)  (PLUS X1 18)))
```

Where X1 is computed from the crop calendar time when at least 50 percent has been planted for wheat. Based on similar considerations, windows W_2 for wheat W_3 and W_4 for barley are defined. These values for X1, et al are determined from models for crop calendars. Essentially they ascertain availability of observations during certain crucial crop development stages. It is an interesting empirical observation that minimum data requirements are different for different geographical regions. This "contextual information" forms a crucial part of the knowledge base. Next, we check to see if the Landsat data are adequate for labeling. For each pixel, there is a list of acquisition dates from which we decide whether or not the region of that pixel is labelable. The following function checks if the list DATA contains an acquisition in particular window by recursive scanning the list DATA.

```
(COND ((NULL DATA) 'ABSENT)

      ((GREATERP (CAR WINDOW)

                 (CAR DATA)
```

(CADR WINDOW)) 'PRESENT)

(T (CHECK WINDOW (CDR DATA)))))

For example, the function (CHECK W1 DATA) will return PRESENT or ABSENT whether or not DATA contains an acquisition in the window W1. This function is evaluated for each of the four windows. Then it is a simple matter to write LISP function which determines whether the minimal data requirements are met, for that particular geographical region. In addition, the knowledge base includes information concerning quality of the data in the form of cloud cover, haze, etc. If such cover exceeds 40 percent, the scene cannot be labeled. The actual labeling procedure or "Paradigm" uses a scatter plot generated from windows 2 and 3 and a decision line which will not be discussed here. It is to be noted that we have given only the outline of the procedure omitting most of the details. An alternative approach, based on an angular statistical measure to crop classification can be found in Ref. [7].

So the knowledge base for an expert system based on the scheme described above can be envisaged to be built around the following skeleton.

1. An eliminating constraint which discards data with excessive haze and/or cloud cover.

2. A production rule which determines the four windows as described earlier. (The window function.)

3. A decision rule which checks the geographical location of the scene against the minimum data requirements in the form of acquisitions in the windows. (The function CHECK.)

4. Data will comprise of crop calendar information as well as Landsat data.

It is interesting to note with the exception of the data, all the rules given above use programs or functions as components in the knowledge base.

Acknowledgments: The author would like to thank his collegues R. P. Heydorn of NASA, R. K. Lennington, E. R. Magness and T. C. Minter of Lockheed and D. Y. Y. Yun of I.B.M.

REFERENCES

[1] R. O. Dudla, P. E. Hart, N. J. Nilsson, and G. L. Sutherland (1978) (Semantic Network Representations in Rule-Based Inference Systems," in Pattern Directed Inference Systems, edited by D. A. Waterman and F. Hayes-Roth

[2] E. A. Feigenbaum (1978) "The Art of Artificial Intelligence – Case Studies in Knowledge Engineering." Proceedings of the National Computer Conference.

[3] R. B. MacDonald and F. G. Hall (1980) "Global Crop Forecasting." Science, Vol. 208, pp 670-679.

[4] A. Newell and H. A. Simon (1972) Human Problem Solving. Prentice Hall, Englewood Cliffs, New Jersey.

[5] W. F. Palmer and E. R. Magness (1981) A Description of the Reformatted Spring Small Grains Labeling Procedure. NASA-Lyndon B. Johnson Space Center Report JSC-16827.

[6] C. V. Srinivasan (1980) Knowledge Based Learning Systems Department of Computer Science. Report DCS-TR-89, Rutgers University, New Brunswick, N.J.

[7] S. G. Wheeler and P. N. Misra (1980) "Crop Classification with Landsat Mutlispectral Scanner Data II." Pattern Recognition, Vol. 12, pp 219-228.

[8] P. H. Winston (1977) Artificial Intelligence, Addison-Wesley, Reading Massachusetts.

[9] P. H. Winston and B. K. P. Horn (1981) "Expert Problem Solving Using IF — THEN Rules," pp. 239-250 in LISP, Addison-Wesley, Reading, Massachusetts.

1830 NASA Road 1
Houston, Texas 77058, U.S.A.

omit

THE UTILIZATION OF SCIENTIFIC COMPUTERS BY THE CASUAL TECHNICAL USER

By Joseph W. Wissel, Ph.D
Lockheed Missiles and Space Company

Technical personnel are confronted almost daily by an infinite variety of scientific computers which have the potential of solving numerous problems. To take advantage of this potential, the individual starts at the man-computer interface. This presentation deals with a small sample of the scientific computer-casual technical user world observed over a two year period.

The People

The casual users population consists of approximately 200 missile system engineers at the Lockheed Missiles and Space Company. None of these engineers are trained programmers. Previous and current experience may involve ownership of a home computer such as an Apple II or TR-3. It may also involve no previous contact combined with an active dislike for any kind of computer. Years of experience at Lockheed may vary from in excess of 25 years to recent college graduates.

The Computers

The computer facilities for the Missile System Engineering Division include two Tektronix graphic computers and peripheral units, and an HP85 desk top computer. There are also other terminals which provide access to main frame computers such as Univac, Dec 20 etc.

The Interaction

The casual user may elect to utilize off the shelf programs available on tape or disc units. Programs available operate on interactive basis. Self teaching programs are also available. The Tektronix and HP computers respond to slightly different forms of basic. All computers have sets of instruction manuals and display error messages for incorrect operations. User acceptance of the computer systems is high. There are varying degrees of programming sophistication exhibited by System Engineering users. Complicated analysis programs and simple scheduling programs have been executed with the scientific computer systems. Experience to date highlights the need for simplification of user operations to assure the user acceptability of future computer systems.

HOW TO TALK TO YOUR COMPUTER AND ENJOY IT

by
T.G. McGinty, R.S. Shirley
The Foxboro Company
Foxboro, Mass.

Abstract

The technology for limited computer recognition of speech is available right now, but actually using it is something else. Unless the human engineering aspects of talking to a computer are fully considered, it will be a very frustrating experience for both the user and the computer. Some lessons learned while using a spoken phrase recognizer are described. They provide some practical guidelines for creating an environment for friendly conversations with your computer.

INTRODUCTION

The capability to have human-like conversations with a computer is not yet available. But a more limited capability, to give spoken commands to a computer and receive verbal responses, is now commercially available and of practical use.

No one has yet seriously attempted to have a computer recognize "normal" speech. The HEARSAY and HARPY projects (13) undertook a substantial part of the task: their goal was to recognize spoken sentences from a vocabulary of less than 100 words, using a slightly constrained grammar and using trained speakers (speakers with good enunciation and no accent). The results are promising, but errors do occur regularly (10%), and the process is slower than real time. The complex computer analysis typically takes much longer than speaking the sentence. The HEARSAY/HARPY work is very informative, but is not yet commercially viable.

Three important factors have become apparent from research in computer recognition of speech:

- the vocabulary
- the grammar
- the speaker

If the vocabulary is limited, clearly the computer has an easier task in determining what has been said. Similarly, if the grammer is well structured, the

computer's task is simplified. Finally, if individual differences between speakers are eliminated, the computer's task is further simplified.

One of the most sophisticated, commercially available connected speech recognizers is the Nippon Electric Company (NEC) speech recognizer (7). The vocabulary for this device is up to 120 words (selectable) and it can process a maximum of 2.4 seconds of continuous input.

Another approach to computer recognition of speech is phrase recognition, an approach taken by such companies as Threshold Technology, Interstate Electronics Corporation (IEC), and Heuristics. The vocabulary for the devices sold by these companies is typically on the order of one hundred words at a time (with easy fast changeover to other sets of one hundred words if desired). The grammar is very limited: only phrases or "utterances" are recognized. Typically utterances can be up to one and a half seconds long, and must be separated by at least one hundred to five hundred milliseconds. The speakers are both constrained and free. The constraint is that each speaker must individually train the speech recognizer by saying every phrase in the vocabulary several (about five) times. The freedom is that the speaker's accent (or even the language) doesn't matter: the pattern of sounds in the utterance is used to form a template against which future utterances will be compared.

For our work we selected the phrase recognition approach. The HEARSAY/HARPY approach is not yet commercially viable and is expensive to implement both in terms of dollars and in terms of manpower. The NEC device was promising, but at the time cost about $60,000. The IEC device we bought was immediately available, and cost $2255.

The aim of our work is human factors engineering. A great deal of thought and software is needed to make computer recognition of speech, even the phrase-recognition type, a practical tool. The bulk of this paper will discuss the human factors lessons learned in our laboratory.

Human Engineering The Phrase Recognizer

Purchasing a phrase recognizer, connecting it to a computer, and applying it to a task is essential, but is not enough. Additional work is necessary to make the computer recognition of speech easy to use: it must be human engineered (see Table I). If the human engineering is not done and done well, the result will be an unhappy user and/or a system which is not used. The best way to understand the human engineering is to go through an example. The problem we chose to solve in order to get practical experience was:

> "Using spoken commands, control the color of a CRT (cathode ray tube) display. Each of two words ("FOXBORO" and "LISTENS") and the background are to have their color independently changed at will to any of eight colors."

We often use the example as a demonstration of voice input and so an attempt was made to make it as self-explanatory and easy to use as possible. Messages to the user are displayed on the bottom line of the screen while commands currently available are given in menu format to the right of the main display area. Only two menues are needed. One menu allows the user to choose which word(s) will have its color changed, and the other presents the user with the color choices.

Since there are several of us that can run the demonstration, the first thing the program does is load a file of patterns created by each of us speaking our name. It then puts the recognizer into recognize mode and waits for someone to say their name. Once a name is spoken the program will load a second template file containing the patterns needed for the speaker to run the demo. The display is put on the CRT, and the speaker can select the word that is to have its color changed.

As simple as the task may seem, its implementation provided many lessons in human engineering. The remainder of this paper will go over the steps shown in Table I, and, along the way, discuss some aspects of the human engineering of talking to a computer.

Select a Recognizer

The reasons for our selection of a phrase recognition device are outlined in the introduction. Basically phrase recognizers were practical, available, and relatively inexpensive. At the time we wanted to purchase a phrase recognizer, the Interstate Electronics Corporation (IEC) recognizer seemed the best for the reasons listed in Table II. Since then, many improvements have come about in computer recognition of speech. Look about, talk with several vendors (see Table III), and choose the recognizer which best meets your needs.

The Interface To Your Computer

Speech recognizers must be connected to your host computer for four reasons: (1) The coded results of your spoken commands must be passed to your computer so it can respond. (2) If multiple sets of phrase templates are to be used (i.e. more than one hundred phrases or more than one speaker are to be recognized), the templates are most easily stored in the host computer or on one of its peripherals. (3) The host computer will support the application (in this case changing colors on a CRT display), and so must be able to control the function of the speech recognizer. (4) Finally, the host computer will be used to run diagnostics on the speech recognizer.

Needless to say, the best situation is one in which the host computer and speech recognizer function as an integrated system. The host computer should respond to voice-entered commands promptly and in such a way that an operator can clearly see that some action has been accomplished.

If large or frequent transfers of voice templates are anticipated, then low speed transmission of data should be avoided. Unnecessary delays will frustrate the operator and make using the system an unpleasant task. When long delays

(more than a second) are unavoidable, the operator should be given positive feedback about what the system is doing. A message like "Please wait while I load the voice templates," goes a long way towards creating a friendly environment.

The speech recognizer we bought is basically a printed circuit (PC) board with components. In addition to interfacing it to the host computer, it was necessary to chassis-mount the PC board in our laboratory, and to supply power to the board. We chose to rack-mount an Intel iSBC 660 chassis, which contains 8 card slots and provides us with adequate expansion capabilities. The chassis contains an Intel 640 power supply which delivers the required 1 amp at +5 volts and 120 mA at + and - 12 volts.

Other suitable mounting arrangements for our device (1) would include the Standard Applied Engineering, Inc. three-slot card cage (Part Number 102498), or the Intel iSBC 604 which provides an interconnection for the 86 pin "Multibus™."

Select a Microphone

The speech recognizer "hears" the spoken phrases through a microphone. The options are numerous. Do you mount the microphone on a wall, on a pedestal, around the speaker's neck, on a headset, or on a panel? Is it best to use a directional microphone, a noise-canceling microphone, or a "regular" microphone? Is it worth using a wireless broadcasting microphone? Should there be an on/off switch?

The above are just a few of the questions that must be answered when selecting a microphone, and they are best answered on an application-by-application basis. Two major criteria for microphone selection are the environment and the task the operator must perform.

Applications for speech recognizers may be found in all types of environments, from the "quiet" office to the noisy machine shop floor. The microphone best suited to one environment may be the worst choice in another, not only from a practical but also from a human factors viewpoint. A noise canceling microphone that eliminates the sound of a nearby typewriter may not be able to eliminate the noise from a grinding machine across the room; perhaps a directional mike is a better solution for eliminating loud continuous noise.

Operator activities when using voice input will also suggest the best type of microphone for the application. If an operator's hands are busy, a continuously active or voice-activated microphone may be necessary. When operator mobility is required, a wireless microphone may be needed. For occasional speech input a push-to-talk switch may be satisfactory; for frequently occurring speech push-to-talk switches will undoubtedly cause much aggravation.

While the application suggests the hardware aspects of microphone selection, software can be used to enhance the actual operator-machine interface. One simple example of this is a voice activated on-off switch. A short phrase (we use "IGNORE-ME") can tell the interface software to ignore all inputs until a specific phrase ("LISTEN-UP") is spoken. This type of switch is especially useful in pre-

venting accidental inputs when the operator is involved in other conversations.

We experimented with three different types of microphones for running our example program. Two were desk-top, push-to-talk mikes which could be locked in the "ON" position, the first was a headset mike with no on-off switch. Our lab is generally "quiet"; as a background noise we have the hum of a nearby PDP 11/60. The only spurious noise is an interoffice paging system, and its volume fluctuates depending on which secretary is paging.

As might be expected, performance was good with each of the microphones. The paging system provided a consistent problem source, especially during vocabulary training. At other times the pages simply resulted in non-recognitions, which eventually became most irritating. The noise cancelling mike was not as sensitive to the paging system as the standard mike, and the headset mike eliminated all but the loudest pages. However, the few errors that we did get always occurred at the most inopportune times. We finally solved the problem by installing a switch so that the paging system could be turned off when the recognizer was in use.

A lesson we did learn about microphones was that placement and mike-to-mouth distance has an effect on performance. We have no guideline for what the mike-to-mouth distance should be, but we know it must be consistent. If a recognizer is trained with the mike held a few inches from the mouth, then the best performance level will be achieved there also. Moving the mike a significant distance further away will result in less than acceptable performance, and, depending on the microphone used, may result in the speaker not being heard at all. An increased voice level can be used to overcome the additional distance, but this will not guarantee better performance.

The best solution when the distance between the speaker and some reference point varies may be to use a microphone that moves with the speaker. We tried this with the headset mike and found the results quite acceptable. The only time a real problem arose was when drinking coffee.

Choose the Phrase to be Recognized

This step is crucial. Not only must words be chosen which are easy and natural to the human for the task at hand, but the resulting phrases must be easily distinguishable by the speech recognizer. Let us consider the human first, then the recognizer.

In his classic article "Words, Words, Words" (2) Alphonse Chapanis states the importance of words for a man/machine interface as follows: "...changes in the words that are used in man-machine systems may produce greater improvements in performance than human engineering changes in the machine itself."

Thus, the words chosen for the human must be natural, easy to remember, and clear in intent. The phrases chosen to control the color display are listed in Table IV.

If the phrases chosen sound alike, no phrase recognizer can distinguish between them. Thus, the phrases "led in" and "leaden," or "write" and "right" cannot be told apart. Other phrases are different, but sufficiently similar as to cause problems, such as "vile" and "file." In order to avoid confusion, it is recommended that you use the speech recognizer and host computer to run some tests. Select the vocabulary with the interests of the human in mind, then train the recognizer. Write a computer program on the host which will compare the templates for each phrase with the templates for every other phrase, and then print out the phrases which have identical or nearly identical templates. This information can then be used to modify the vocabulary until all the phrases can be uniquely distinguished by the speech recognizer.

Of course, the difficult part of the above technique is defining what constitutes an "identical" template. Our first (and most successful) attempt at measuring closeness of templates was to compare corresponding bits in the templates and count the occurrence of equal bits. This measure is reported as a percentage matched (number of corresponding, equal bits ÷ total number bits in a template). We found that whenever a comparison of two templates resulted in a percentage match of 80% recognition errors would occur. As the percentage match measure increased, so did the frequency of recognition errors for the words involved.

What we are really measuring here is the similarity of sounds (words) from the recognizer's point of view. When two words sound alike (high percentage match measure) some action should be taken to prevent recognition errors from occurring. Either additional training of the problem words or replacing one of the words in the vocabulary should solve the problem. In either case, retesting of the vocabulary is necessary to insure that no additional similarities have been created.

Select the Refusal/Error Tradeoff

There are three parameters to consider with respect to error rates in computer recognition of spoken phrases:

N The number of spoken phrases

e The number of errors in recognizing phrases (i.e., mistakes)

r The number of rejects in recognizing phrases (i.e., no recognition)

Using these parameters, two commonly used definitions of error rate, E, are

$$E_1 = \frac{e}{N-r} * 100 \qquad (1)$$

$$E_2 = \frac{e+r}{N} * 100 \qquad (2)$$

and the recognition rate, R, is defined (in percent) as

$$R = \frac{N-e-r}{N} * 100 \qquad (3)$$

The use of E_1 emphasizes the avoidance of mistakes in recognition. The use of E_2 considers both mistakes and lack of recognition as equally bad. E_1 is more commonly used; it is generally considered best for the computer to occasionally ask the operator to repeat a phrase than for the computer to make a mistake. The use of R states that what the operator really wants is 100% recognition with no errors at all.

There is a basic tradeoff between e, the number recognition errors, and r, the number of non-recognitions. A phrase recognizer has a tolerance factor which is used to check for recognition of a phrase. When a phrase is spoken, it is analyzed and a temporary template created. This template is then compared to the pre-programmed, stored templates for all the allowable phrases. The closest stored template is then tentatively considered a match. The final step, before declaring a match, is to see how close the temporary template is to the closest stored template. If the allowable tolerance is small, there will be very few recognition errors, but more non-recognitions. If the allowable tolerance is large, there will be very few non-recognitions, but more recognition errors. A very large tolerance will lead to the selection of the closest phrase in the vocabulary, regardless of what is said.

Fortunately, vendors of speech recognizers can provide some guidelines for determining what the tolerance factors for their devices should be. These guidelines provide a starting point from which the best tolerance factor can be found for each speaker for each application.

Determine the Computer Responses

Even a fresh kid is not as bad as a fresh computer. Similarly a computer which is wordy, boring, repetitive, confusing or in error is generally worse than a human with the same faults. Humans wave their hands, make faces, and respond to confused looks, and so can cover up for many of their faults. The computer has to be right, or at least provide a channel for recovery via more communication.

What words and/or actions will the computer use for each and every response? If the words are too long, they will soon make the operator impatient. If they are ambigous, they will lead to poor communication and error. If they are too polite or impolite, they will lead to bad feelings within the human.

As users of speech recognition systems we should not limit our efforts to design of input vocabularies. Output (response) vocabularies are equally important. Any word, phrase, or action we choose for a computer response should be clear, concise, unambiguous, and most of all, it should be friendly. Consider the possible responses when the host computer detects a non-recognition. An error message like "NOT RECOGNIZED," or "WHAT?," or even "SORRY, PLEASE SAY IT AGAIN" might be

displayed. All of these are concise, each is to some degree informative, and all are equally easy to output. Why choose "NOT RECOGNIZED"? It states the problem, but doesn't really tell the user how to proceed. "What?" implies there is a problem and suggests the next action to the user, but does so in an impersonal and hostile way. The third choice admits there is a problem, tells the user how to proceed, is polite, and, in fact, appears friendly. This type of response should be the first choice of any human factors engineer.

An even more impersonal response by a computer is the single character prompt used by many operating systems and imitated by many programmers in theirown software. When words are being used to communicate with a machine, the machine should use words to communicate as well. After calling up our example program on our PDP 11/60, we display "WE ARE LISTENING" rather than use the single character "*" as so many other programs do. So far no one has had any difficulty in deciding what to do next. It is obvious that a voice command is in order.

We have tried a novel technique for operator feedback that involves setting up several equivalent messages for each response. Whenever a response of a particular type is needed, a random number generator is used to select a message. The fact that there are several possible messages that mean the same thing does not seem to bother the users. On the contrary, they may even be paying more attention. The boring, repetitive and wordy aspects of computer response have been partly eliminated, and a friendly user environment has been created.

Choose the Computer Response Modes

There are many possible ways in which a computer can respond, with two obvious ones being words on a CRT and/or spoken words on a speech synthesizer. There is preliminary indication that if a human is speaking to a computer, receiving spoken words in return is comfortable, though simultaneous printed words (on the CRT) is probably best.

The most important consideration here is that the computer does respond, both quickly and in an unambiguous way. Only one thing is more frustrating to a user than waiting for something to happen: waiting and having nothing happen. Each word spoken by a user should have some immediate, positive response. Something as simple as displaying the word just heard will serve to reinforce the user's confidence in himself and in the system being used.

Program Your Application

Having made all the decisions, the final step is the implementation of your application on the host computer.

In a general sense, integrating a speech recognizer into an applications system is not much different than integrating any other peripheral. The recognizer has a finite set of functions which it can perform, a set of commands for invoking those functions, and a fixed response for each function it executes. Taking this

somewhat simplistic view reduces the recognizer to the status of something like a terminal.

The primary function of the recognizer is to tell a host computer what it thinks it "heard," usually by outputting a numeric code corresponding to a vocabulary entry. This is directly analogous to a keyboard controller echoing what a user has typed. Once the host computer has requested and received an input, it can take some appropriate action, provide the user with any necessary feedback, and await the next input. The recognizer, when in recognition mode, is quite easy to deal with.

All other functions of a recognizer can be considered as supportive of the recognition function. No doubt this is why a system/applications designer must spend a significant amount of time determing how these functions are best implemented. Should the user have the ability to set tolerance factors? Where are vocabulary templates stored? How many users should have access to the system via voice input? Are template update and save functions really necessary? Should user vocabulary training and updating be incorporated into the application or should these functions take place off-line? These are only a few of the questions a designer must answer.

Often the application itself will provide an indication as to what functions should be included. The worst mistake a designer can make is to ignore the needs of the application. If an application is suited to speech recognition, the overall design and integration will be a pleasant, easy task. Force-fitting a recognizer into a system will be difficult, result in poor design decisions being made, and usually the end product and its user will suffer.

Conclusions

Talking to your computer can be an enjoyable, productive experience. Accomplishing this depends as much on the human engineering of speech recognition as it does on the available technology. Proper attention must be given to the microphone selection, vocabulary, recognition error trade-offs, and desired computer responses before integrating a speech recognizer into your application.

Table I

STEPS FOR HUMAN ENGINEERING THE COMPUTER

Recognition of Speech Using a Phrase Recognizer

Select the Recognizer

Interface It To Your Computer

Select a Microphone

Choose the Phrases To Be Recognized

Select The Refusal/Error Tradeoff

Determine The Computer Responses

Choose The Computer Response Modes

Program Your Application

Table II

CHARACTERISTICS OF THE INTERSTATE ELECTRONICS CORPORATION VOICE RECOGNITION MODULE 102

- Single Board
- Parallel and Serial Interface Ports
- ASCII Characters as Data
- Serial Baud Rates Selectable to 9600 Baud
- Stand-Alone or Host-Computer Modes
- Large (100) Word Vocabulary
- Reject Threshold Parameter Control
- "Common" Vocabulary Parameter, allowing a portion of the vocabulary to always be active
- "Selectable Syntax Structure" via the ability to limit the section of the vocabulary to be considered during recognition

Other vendors offer speech recognizers (see Table III). Choose the one best suited to your needs.

Table III

*Vendors of Speech Recognition Devices

Auricle, Inc.
Centigram Corp.
Dialog Systems, Inc.
Heuristics, Inc.
Interstate Electronics Corporation
Nippon Electric Co. (NEC)
Threshold Technology, Inc.

*Other vendors may exist - speech recognition is a very dynamic technology.

Table IV

The Phrases Used to Control the Color of the CRT Display

State 1:

Foxboro
Listens
Both
Background
Ignore Me
Listen Up

State 2:

No Change
Black
Blue
Cyan
Green
Purple
Red
White
Yellow

References

1. -------, "Voice Recognition Module Reference Manual," Interstate Electronics Corporation, 1980.

2. Chapanis, A., "Words, Words, Words," Eighth Annual Meeting of the Human Factors Society, Presidential Address, October, 1964.

3. Denes, P.B.; Pison, E.N., "The Speech Chain: The Physics and Biology of Spoken Language," Anchor Book, 1973.

4. Ferber, L., Yilmaz, H., Kellett, H., and Doohovskoy, A., "Continuously Automatic Speaker Adaptation," AD-A082 620/6 Rome Air Development Center, April 1980.

5. Kolbus, D.L., "Computer Speech Communication, "SRI International Business Intelligence Program Report, 1979.

6. Lea, W.E., Ed., "Trends in Speech Recognition," Prentice-Hall Signal Processing Series, 1980

7. Schmandt, C., "The Intelligent Ear, A Graphical Interface to Digital Audio," Architecture Machine Group Report 224, M.I.T., 1980.

8. Shochet, E., and Lemanski, R., "Automated Flight Plan Filing By Simulated Voice Recognition, "FAA-RD-80-56, FAA-CT-80-15, Federal Aviation Administration Technical Center, August 1980.

9. Simmons, E.J. Jr., "Speech Recognition Technology," Computer Design, June 1979, pp. 95-101.

10. Reddy, D.R., "Eyes and Ears for Computers," AD-760 153, Air Force Office of Scientific Research, National Science Foundation, March 1973.

11. Warren, C., Teja, E.R., "Computers and Peripherals," EDN, Vol. 25, No. 22, December 1980, pp. 310-324.

12. Welch, J.R., "Automatic Speech Recognition - Putting it to Work in Industry," IEEE Computer, Vol. 13, No. 5, May 1980, pp. 65-73.